ADVANCES AND APPLICATIONS OF STATIC AND DYNAMIC CORRELATION METHODS IN AB-INITIO QUANTUM CHEMISTRY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Debashree Ghosh

February 2010

ADVANCES AND APPLICATIONS OF STATIC AND DYNAMIC CORRELATION METHODS IN AB-INITIO QUANTUM CHEMISTRY

Debashree Ghosh, Ph.D.

Cornell University 2010

Computational chemistry has made remarkable progress in the last couple of decades due to the availability of better and more powerful computers. However, there is definitely more room for improvement. There are problems in material science and biology that require 10^2 to 10^4 atoms to be considered and the importance of such systems cannot be over-emphasized. Addressing such systems with *ab-initio* methods is still a quantum chemist's dream.

Brute force solutions of the Schrödinger equation are not possible for reasonably interesting and important systems due to the large number of Slater determinants that have to be stored in a computer's memory. Thus electronic structure theorists have made chemically and mathematically intuitive approximations to simplify the problems.

A particularly difficult class of systems that is very interesting to the theoretical and experimental chemists is the strongly correlated systems. High T_c superconductors, nanotubes, graphene sheets, transition metal complexes and photosynthetic materials all fall under this category. These systems are too large to be treated by traditional quantum chemical methods. Another reason why such systems have not been studied sufficiently by *ab-initio* method is that they comprise of complicated strong electron-electron correlation (and therefore the name strongly correlated), thus, making them impossible to be qualitatively defined by a molecular orbital picture, which has been a quantum chemist's fa-

vorite tool for a long time now.

Therefore, in this thesis, we have tried to develop, improve and apply methods for treating such strongly correlated systems. The thesis is broadly divided into two parts. In the first part, we discuss the orbital optimized DMRG method developed to treat static correlation in large strongly correlated systems. In the next part we discuss the two methods that can treat dynamic correlation in such complicated systems.

BIOGRAPHICAL SKETCH

Debashree Ghosh was born in Kolkata, India on 7th September, 1981 as the eldest daughter of Debabrata Ghosh and Sudipta Ghosh. Her formative years were spent in a nearby town, Durgapur where she completed her primary education from St. Michael's School. Later on she moved with her parents to her ancestral home in Serampore where she completed her secondary education in St. Joseph's Convent, Chandannagar in 1997 and her higher secondary education in Serampore College in 1999.

Debashree joined Presidency College, Kolkata for her Bachelor of Science degree with Chemistry major and Physics and Mathematics minors. It was here that she was taught by a lot of motivated teachers who inspired her to go for higher education and research. Graduating from Presidency College in 2002, she joined Indian Institute of Science, Bangalore in Integrated Ph.D. program. Debashree left Indian Institute of Science after her Masters in Chemistry and joined Cornell University for Ph.D.

Dedicated to my grandfather, who would have been very proud.

ACKNOWLEDGEMENTS

I would like to thank all the people who helped me during my doctoral studies.

I especially want to thank my adviser, Prof. Garnet Chan for his guidance during my research and study at Cornell University. His insight, quest for perfection and passion for science was a constant inspiration to me.

I also would like to express my gratitude towards my other committee members, Prof. Hoffmann and Prof. Ezra, and all the professors in Cornell University who taught me. I express my gratitude for Prof. Takeshi Yanai, my mentor who taught me all about electronic structure theory.

I would like to thank Dr. Dominika Zgid for not only academic help, but also being a good friend. Special thanks to Jon for being a good friend and a great lab mate. I express my gratitude for all my lab mates (Johannes, Haitao, Eric, Claire, Hitesh, Jesse) for creating such a nice work environment.

All my friends in Ithaca deserve a very special thanks. I could not have been able to live here and finish my work without the endless support from Snad, Soumya and Ananda. Thanks to little Debashree for being a very good roommate and friend, and to Anupam for a lot of chitchat. Thanks to all my batchmates, especially Scott, Paulina, Nesha, Ruchi, Sumana, Amrita, Ananda and Ankush. Special thanks to Scott, Jon and Eric for a lot of ping-pong.

I would also like to thank Aleya, Sreya, Maha, Ishita, Debarshini and Devasena for all their support. Thanks to Akash and Sutirthada for calling me up so many times only to hear me complain.

My deepest gratitude goes towards my family for their love and support. Special thanks to my sister. Last but not the least I would like to thank my husband, Suman for all the love and understanding and immense patience with me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Strong and weak correlation

In quantum chemistry and more specifically in electronic structure theory, the main aim of a researcher is to solve the Schrödinger equation, $H\psi = E\psi$ associated with the quantum chemical Hamiltonian, formed by applying the Born-Oppenheimer approximation to the full Hamiltonian for the electrons and nuclei in the molecule,

$$[-\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 - \sum_A \frac{\hbar^2}{2M_A}\nabla_A^2 - \sum_{A,i}\frac{Z_A e^2}{4\pi\epsilon_0 r_{Ai}} + \sum_{A>B}\frac{Z_A Z_B e^2}{4\pi\epsilon_0 R_{AB}} + \sum_{i>j}\frac{e^2}{4\pi\epsilon_0 r_{ij}}]\Psi(\mathbf{r};\mathbf{R}) = E\Psi(\mathbf{r};\mathbf{R}). \quad (1.1)$$

Eqn. 1.1 is the full time independent Schrödinger equation, solving which we can obtain the combined nuclear and electronic wavefunction of a molecule, where **r** and **R** denote the electronic and nuclear coordinates and the indices i and j refer to electrons and A refers to nuclei. To make the equation more compact, we introduce the atomic units, $\hbar = |e| = \frac{1}{4\pi\epsilon_0} = m_e = 1$. Incorporating the atomic units in Eqn. 1.1 we get,

$$[-\frac{1}{2}\sum_i \nabla_i^2 - \sum_A \frac{1}{2M_A}\nabla_A^2 - \sum_{A,i}\frac{Z_A}{r_{Ai}} + \sum_{A>B}\frac{Z_A Z_B}{R_{AB}} + \sum_{i>j}\frac{1}{r_{ij}}]\Psi(\mathbf{r};\mathbf{R}) = E\Psi(\mathbf{r};\mathbf{R}). \quad (1.2)$$

The left hand side of the equation consists of a kinetic part due to the electronic motion $\hat{T}_e(\mathbf{r}) = -\frac{1}{2M_A}\sum_i \nabla_i^2$, a nuclear kinetic term $\hat{T}_N(\mathbf{R}) = -\frac{1}{2}\sum_A \nabla_A^2$, the electron-nuclear attractive potential term $\hat{V}_{eN}(\mathbf{r};\mathbf{R}) = -\sum_{A,i}\frac{Z_A}{r_A i}$, the nuclear

nuclear repulsion term $\hat{V}_{NN}(\mathbf{R}) = \sum_{A>B} \frac{Z_A Z_B}{R_{AB}}$ and the electron-electron repulsion term $\hat{V}_{ee}(\mathbf{r}) = \sum_{i>j} \frac{1}{r_{ij}}$ respectively.

After applying the Born-Oppenheimer approximation, we can obtain a residual Schrödinger equation for the electrons only,

$$[-\frac{1}{2}\sum_i \nabla_i^2 - \sum_{A,i} \frac{Z_A}{r_{Ai}} + \sum_{A>B} \frac{Z_A Z_B}{R_{AB}} + \sum_{i>j} \frac{1}{r_{ij}}]\Psi(\mathbf{r}) = E_{el}\Psi(\mathbf{r}),$$

$$[\hat{T}_e(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r};\mathbf{R}) + \hat{V}_{NN}(\mathbf{R}) + \hat{V}_{ee}(\mathbf{r})]\Psi(\mathbf{r}) = E_{el}\Psi(\mathbf{r}). \tag{1.3}$$

To understand the challenges involved in electronic structure problem, it is customary to start with the simplest possible approximation to its solution, the Hartree-Fock (HF) approximation [1–4]. In this approximation, one treats the electrons as moving in a *mean field* of the other electrons, i.e. the electrons ignore the effect of the other electrons except in a mean field kind of way.

Going beyond the mean field approximation to consider a picture where the electrons feel the presence of other electrons individually, involves considering the *electron correlation*. Electron correlation, as defined by Löwdin [5], is the difference between the mean field or Hartree-Fock approximation and the exact solution of the full *non-relativistic* Schrödinger equation.

According to the degree of electron correlation, all quantum chemical systems can be classified into weakly correlated and strongly correlated. The term *weakly correlated* refers to the systems that can be qualitatively well described by simple one electron theories like density functional theory (DFT) [6–8] or HF theory. Although HF theory appears to be a drastic approximation, it has served well as a starting point for most of these systems. In order to calculate experimental observables with sufficient accuracy, there are well-known and well-

developed post Hartree-Fock methods like configuration interaction (CI) [9–12], Møller Plesset (MP) perturbation [13] and coupled cluster (CC) [14, 15] theories.

The term *strongly correlated* refers to materials or molecules that are not qualitatively described by simple one-electron theories. Among this class of compounds are interesting materials that show unusual electronic and magnetic properties, most well known being high T_c superconductors, transition metal complexes, conjugated polymers and biological compounds of photosynthetic importance. In essence, not all properties of these systems can be qualitatively understood from a simple free-electron model, nor are they fully ionic in character, thus making these systems very interesting as well as difficult to understand comprehensively. These systems have near-degenerate valence orbitals and the Coulomb repulsion becomes comparable to the orbital energy spacing. Therefore, it is impossible to pinpoint a single valence electronic configuration as the predominant configuration. Thus, it seems logical that Hartree-Fock, being a single determinant approximation is an inadequate starting point.

As a simple example, if we take two H atoms in the dissociation limit , both the determinants $|\sigma_{g\uparrow}\sigma_{g\downarrow}\rangle$ and $|\sigma_{u\uparrow}\sigma_{u\downarrow}\rangle$ are equally important since $|\sigma_g\rangle$ and $|\sigma_u\rangle$, i.e. the bonding and antibonding orbitals are degenerate at the limit of no interaction. Thus, there is need for at least two closed shell Slater determinants $[\sigma_g(r_1)\sigma_g(r_2) + \sigma_u(r_1)\sigma_u(r_2)]$ of equal weight to describe the H_2 molecule at this limit. If we start with any one of the determinants and try to add in correlation, we are treating the two determinants differently which should have been treated similarly a priori. The *correct* dissociation of H_2 is into H radicals $\mathrm{H}\cdot\mathrm{H}\cdot$ as expected intuitively. However HF theory cannot treat more than one Slater determinant and, therefore predicts equal probability of dissociating a H_2 molecule

into H^+H^- or $H \cdot H\cdot$. For this reason the restricted Hartree-Fock energy of the hydrogen molecule at very large separation is unphysically high. Another example where HF theory fails spectacularly is the Cr_2 dimer. Here, HF theory fails to predict any binding. For conjugated polyenes, it is unable to predict the presence of low-lying excited states that are formed by more than single electron excitations. Thus, we can come to the conclusion that the single electronic configuration picture is not adequate to understand situations like bond breaking, e.g. the H_2, N_2 and Cr_2 bond breaking, and excitations in conjugated systems, e.g. organic polymers and biologically important systems, as well as metal complexes of material and biological importance.

In these situations, more than one electronic configuration needs to be treated on an equal footing. The name *multireference* is used for such quantum chemical methods that have the capability to treat large number of valence configurations. Examples of such methods are complete active space configuration interaction (CASCI), complete active space self-consistent field (CASSCF) [16–18], complete active space perturbation method (CASPT2) [19–21], multireference Møller Plesset perturbation theory (MRMP)[22] and multireference configuration interaction (MRCI) [23, 24]. A more recently developed multireference method is density matrix renormalization group (DMRG) [25].

Here, it should be pointed out that, of course, if we could do a full configuration interaction (FCI) calculation on any system, i.e. consider all possible electron configurations as

$$|\Psi\rangle = \sum_{n_1 n_2 n_3 \ldots} \Psi^{n_1 n_2 n_3 \cdots} |n_1 n_2 n_3 \ldots n_k\rangle, \tag{1.4}$$

where n_i spans over $|-\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle$, we would be able to obtain the *correct* (within the basis set) answer to all our questions. However in Eqn. 1.4, the basis

size scales as 4^k. The exponential scaling makes the FCI calculation feasible for at most 16 electrons in 16 orbitals with the fastest and newest computers. This means that all relevant molecules with sufficiently large basis sets are too big to do FCI on and therefore there is the need to get approximate theories that do not compromise too much of the accuracy but which are more tractable than FCI.

The main challenge in formulating a method that can handle strongly correlated systems is the need to treat different and a priori unknown electronic configurations at an equal footing since it is difficult to devise a method which does not simply consider *all* the configurations within a set of valence or active orbitals. In fact, this is the strategy taken by most of the multireference methods that were listed above which makes them scale exponentially with respect to the valence or active orbitals and thus, the system size in most cases.

However, DMRG uses a different approach than the other multireference methods. It circumvents this implicit exponential scaling with the system size. DMRG has a polynomial cost scaling, which will be explained in the later chapters. Therefore, we are interested in trying to develop DMRG based multireference methods.

1.2 Static and dynamic correlation

In this section, we will try to introduce some terminology used in electronic structure theory. Firstly, it is customary to divide the orbitals broadly into core, active and virtual orbitals. *Core orbitals* are the low energy orbitals that are always largely occupied and are not involved directly in chemical reactions. *Virtual orbitals* are the very high energy orbitals that are almost never occupied and

also do not play a major role in chemical reactions. The core and virtual orbitals together form the non-valence or external orbitals. *Active orbitals* are the orbitals that are in between the core and virtual orbitals in energy and are generally made up of the complete space of valence orbitals and are the most important orbitals that play a vital role in chemical reactions and bond breaking.

Static correlation is the correlation arising from near degeneracies in the valence or active orbitals and associated with the multireference nature of the system. It can also be understood as synonymous with strong correlation. Since, static correlation comprises of the correlation arising from the valence orbitals, this can give a qualitatively correct starting point for most systems. *Dynamic correlation* arises from correlations to and from non-valence or external orbitals. Due to the cusp like nature of the wavefunction at the electron-electron coalescence, the dynamic correlation is slowly convergent with the number of external orbitals. Fortunately, in most cases the contribution of dynamic correlation to the energy difference is much smaller than the contribution to the total energy. While the static correlation alone gives the correct qualitative description in most problems, dynamic correlation has to be considered for quantitative accuracy needed for the calculation of experimentally observed properties.

We can divide quantum chemical methods into those that can treat static and dynamic correlation. For the treatment of static correlation, there are methods like CASSCF, CASCI and DMRG. The dynamic correlation is incorporated as a correction on top of this static correlated reference wavefunction $|\Psi_0\rangle$. The exponential ansatz,

$$|\Psi\rangle = e^T|\Psi_0\rangle \tag{1.5}$$

where T is an excitation operator from the active to external space, has been

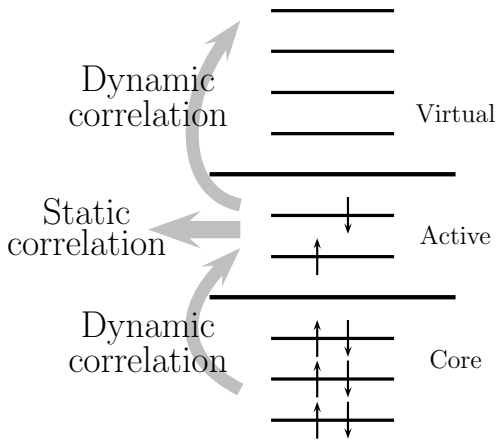


Figure 1.1: Dynamic and static correlation: dynamic correlation is the correlation out of the active space while static correlation is the correlation arising from the active space orbitals.

known to provide an accurate and economical solution to the dynamic correlation problem. In the case of systems with single reference character, experimentally observed properties can be quite accurately calculated by coupled cluster (CC) theory, which corresponds to the exponential ansatz in Eqn. 1.5, where $|\Psi_0\rangle$ is a single determinant. However for multireference problems, the problem of incorporation dynamic correlation is still much more open and a completely satisfactory analogue of CC theory has yet to be found. When the reference wavefunction is a multireference state, some of the natural simplification and advantages of the exponential ansatz are lost, thereby making the CC equations a lot more complicated and impractical for most cases.

In Table. 1.1, dynamic correlation methods for both the single reference and multireference systems are tabulated.

Table 1.1: Hierarchy of quantum chemical methods for dynamic correlation. The method bracketed in the table is explained in this thesis.

Dynamic correlation for single reference	Dynamic correlation for multireference
MP2	CASPT2
CISD and variants	MRCI and variants
CCSD and variants	(Canonical transformation)

In this thesis, we describe some work to target both static and dynamic correlation in multireference systems. The method we use to target the static correlation problem is orbital optimized density matrix renormalization group [26] and the methods we use to target the dynamic correlation problem are cumulant approximated n-electron valence perturbation theory (cu-NEVPT2) [27] and canonical transformation theory [28].

1.3 Scope of the thesis

In Chapter II we start with introducing the concepts of the density matrix renormalization group and explaining the algorithm and the properties of the method. Chapter III deals with the theory and implementation of orbital optimization with density matrix renormalization group that improves the applicability of DMRG to static correlation in highly multireference systems with nontrivial active spaces. In Chapter IV we describe the application of the above mentioned method to conjugated polyenes and carotenoids. This consists of the static correlation part of my thesis. In the dynamic correlation part, start-

ing with Chapter V, we explain the theory and implementation of a cumulant approximated n-electron valence perturbation theory that can be used for large systems. In Chapter VI, we describe the theory and application of canonical transformation theory to small polyenes. Chapter VII brings together the static and dynamic correlation methods with the joint application of orbital optimized DMRG and cu-NEVPT2/canonical transformation theory to carotenoids. We conclude in Chapter VIII with a discussion on computing the three particle reduced density matrix from a DMRG wavefunction and other future directions.

BIBLIOGRAPHY

[1] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).

[2] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications, Inc., 1989.

[3] R. McWeeny, *Methods of Molecular Quantum Mechanics*, Academic Press Inc., 2nd edition, 1996.

[4] D. B. Cook, *Handbook of Computational Quantum Chemistry*, Oxford University Press, 1998.

[5] P.-O. Löwdin, Phys. Rev. **97**, 1509 (1955).

[6] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[7] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[8] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, International Series of Monographs on Chemistry 16, Oxford Science Publication, Oxford, United Kingdom, 1989.

[9] S. F. Boys, Proc. Roy. Soc. London. **A201**, 125 (1950).

[10] R. K. Nesbet, Proc. Roy. Soc. London. **A230**, 312 (1955).

[11] B. O. Roos, Chem. Phys. Lett. **15**, 153 (1972).

[12] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory*, John Wiley & Sons Inc., 2000.

[13] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[14] J. Cizek, J. Chem. Phys. **45**, 4256 (1966).

[15] G. D. Purvis and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

[16] D. R. Hartree, W. Hartree, and B. Swirles, Phil. Trans. Roy. Soc. (London) **A238**, 229 (1939).

[17] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, Chem. Phys. **48**, 157 (1980).

[18] B. O. Roos, The complete active space self-consistent field method and its applications in electronic structure calculation, in *Advances in Chemical Physics; Ab Initio Methods in Quantum Chemistry II*, edited by Wiley, page 399, Chichester, 1987.

[19] K. Andersson, P.-Å.. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).

[20] K. Andersson, P.-Å.. Malmqvist, and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).

[21] K. Hirao, *Recent Advances in Multireference Methods*, volume 4 of *Recent Advances in Computational Chemistry*, World Scientific Publishing Co. Pte. Ltd., Singapore, 1999.

[22] K. Hirao, Chem. Phys. Lett. **190**, 374 (1992).

[23] H. F. Schaefer III and F. E. Harris, Phys. Rev. Lett. **21**, 1561 (1968).

[24] R. J. Buenker, S. D. Peyerimhoff, and W. Butscher, Mol. Phys. **35**, 771 (1978).

[25] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[26] D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, J. Chem. Phys. **128**, 144117 (2008).

[27] D. Zgid, D. Ghosh, E. Neuscamman, and G. K.-L. Chan, J. Chem. Phys. **130**, 194107 (2009).

[28] D. Ghosh, E. Neuscamman, T. Yanai, and G. K.-L. Chan, J. Chem. Theor. Comput. (submitted).

CHAPTER 2

DENSITY MATRIX RENORMALIZATION GROUP

2.1 Need for renormalization

The renormalization group (RG) [1] has been used to study phase transitions and critical phenomena for a long time. The traditional need for renormalization arises when there is no distinct separation between length or time scales. For example, at a critical point, there are very large macroscopic fluctuations along with fluctuations at all intermediate length scales. In such drastic situations, the idea of the renormalization group is used to reduce the number of degrees of freedom across all the length scales into something tractable.

In quantum chemistry, the Hamiltonian is written in second quantization as,

$$H = \sum_{ij} t_{ij} a_i^\dagger a_j + \sum_{ijkl} v_{ijkl} a_i^\dagger a_j^\dagger a_k a_l. \tag{2.1}$$

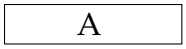
While from this form, it looks like there are only one and two body interactions, the motion of one electron can be relayed to all other electrons through a domino effect to a greater or lesser degree. This domino effect can be efficiently described in a renormalization group framework, even when we are away from criticality, and this is the origin of the use of RG in quantum chemistry.

2.2 Numerical renormalization group

The numerical renormalization group (NRG) technique was developed by Wilson to solve the Kondo problem [2, 3]. Shortly after this, there was consider-

able interest in applying this technique to other related problems. Quantum lattice models such as the Heisenberg and Hubbard models in one dimension appeared to be ideal candidates for the application of this method. However, this approach proved to be rather unreliable, particularly when compared to existing methods.

In traditional RG, we consider some quantum system 

described by the quantum degrees of freedom $\{A\}$ expressed on a subset of quantum sites on a finite lattice. During the RG, we imagine growing ("blocking") 

to cover additional sites on the lattice, while truncating ("decimating") the degrees of freedom at each step. The main steps in the numerical renormalization group method are therefore:

1. From two simple blocks A build a compound block AA and construct the corresponding Hamiltonian H_{AA}.
2. Diagonalize H_{AA} and obtain the m lowest eigenvectors U.
3. From U obtain a subset $\bar{U}$ of m eigenvectors with the lowest eigenvalues. Transform the Hamiltonian to this renormalized basis $\bar{U}$. This defines a truncated (or "decimated") representation A' of the compound block AA, such that e.g. the Hamiltonian becomes $H_{A'} = \bar{U}^\dagger H_{AA} \bar{U}$.
4. Define A' to be the new "simple" block A.
5. Iterate from step 1.

2.2.1 Problems of numerical renormalization group

The main source of error in using numerical RG for spin problems comes from using the lowest energy eigenstates of H_{AA} as the decimated representation. Since H_{AA} has no connection to the rest of the lattice, its eigenstates may have unwanted characteristics, e.g. unphysical nodes at the ends of the composite block AA which correspond to the middle of the complete lattice.

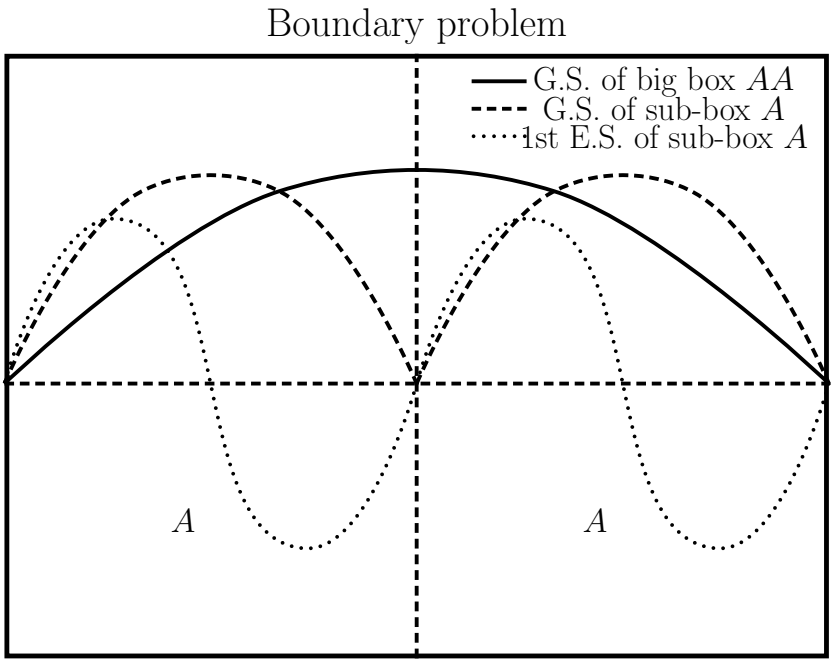


Figure 2.1: The ground state of a particle in a box made from the ground and excited states of sub-boxes. The dashed lines denote the ground and first excited state of the sub-boxes A, while the solid line is the correct ground state of the composite box AA.

This boundary value problem can be understood by a very simple example. In a particle in a box, Fig. 2.1 we can see that if we start with a small box (shown by the dashed lines), and use the lowest energy eigenstates for this building block

to solve for ground state wavefunction of the composite box AA, we are always faced with a node in the middle which is unphysical in the composite box but which occurs due to the boundary condition of the building blocks. No matter how many states we keep for the small box, we cannot remove this unphysical node. This problem is the reason why numerical RG was later reformulated using a density matrix of the composite lattice to determine the decimation procedure.

2.3 Density matrix renormalization group

The density matrix renormalization group (DMRG) was developed by White and Noack [4, 5] as a numerical technique for finding accurate approximations to the ground state and low-lying excited states of strongly interacting quantum lattice models such as the Heisenberg and Hubbard model especially in one dimension.

The main steps in the density matrix renormalization group are [6–8]:

1. A site commonly called a dot is added to the initial block A to form the superblock A'.
2. The complete Hamiltonian of the system is constructed using $H_{A'} + H_{B'}$ where B' is the rest of the lattice other than A' and the Hamiltonian of the B' part is a guess Hamiltonian to start with. Here it should be noted that A is commonly referred to as the system block and the B as the environment block and in between the blocks there are two sites or dots.
3. Now, that we have the approximate Hamiltonian for the complete system

$H_{A'B'}$, we solve for a few low-lying eigenstates of the Hamiltonian using the Lanczos or Davidson algorithm.

4. The density matrix of the system block is formed by tracing out the environment from the eigenstates that we have solved for.
5. The density matrix is diagonalized.
6. Using M most significant eigenvectors of the density matrix, we project the space $\{A'\}$ to a smaller space of size M.
7. We replace A with the decimated block A' and iterate from step 1.

Thus, we see that DMRG differs from NRG in two main ways:

1. The approximate Hamiltonian of the full system, and not just the superblock, is used to decide on the degrees of freedom that are retained. This fixes the problem of unphysical boundary conditions.
2. The criterion for keeping the degrees of freedom is the largest weight eigenstates of the density matrix and not the lowest eigenstates of the Hamiltonian.

2.4 DMRG algorithm in quantum chemistry

The DMRG was applied in quantum chemical problems for the first time by Fano *et al*. [9] where the cyclic polyenes were studied using Pariser-Parr-Pople Hamiltonian. In 1999, the first application of the DMRG with *ab-initio* Hamiltonian to H_2O molecule was presented by White [10]. Later there were several other applications and implementations by various groups such as those of Mitrushenkov, Chan, Legeza [11–15].

As mentioned earlier we are trying to solve for the eigenstates of the quantum chemical Hamiltonian given in Eqn. 2.1. We first imagine that the orbitals that are correlated are ordered in some way as sites on a lattice, so as to make connection to the original RG formulation. We know that from the simple particle in a box example, we need to not only consider a system block that is growing in the renormalization procedure but also the remaining sites in the lattice, or environment block. At any stage in the RG algorithm or "sweep", therefore, there are two blocks L and R (corresponding to system and environment) which vary in size along the sweep and two smaller blocks $\bullet_L$ and $\bullet_R$ ("dots") which are used to enlarge the L and/or R blocks in the blocking procedure, as shown in Fig. 2.2. When sweeping from left to right, L is the system and R is the environment, while in the reverse direction, their roles are also reversed.

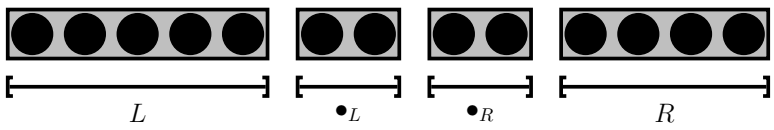


Figure 2.2: The standard block configuration. The lattice is divided into left block L, left dot $\bullet_L$, right dot $\bullet_R$ and right block R. $\bullet_L$ site is blocked with L and $\bullet_R$ is blocked with R.

Sweep iterations are divided into three major steps:

1. Blocking

   Each of the blocks L and R spans a set of orbitals (or "sites") and consists of a subset of the corresponding Fock space (as we have truncated the Fock space using RG along the sweep). In the blocking step, the L and R blocks are combined with their adjacent sites to make a superblock of the system

and environment. That is block L containing states $\{l\}$ are blocked with $\bullet_L$ giving the superblock consisting of $\{l\} \otimes \{\bullet_l\}$ states where $\{l\} \in F_l$ and $\{\bullet_l\} = F_{\bullet_l}$.

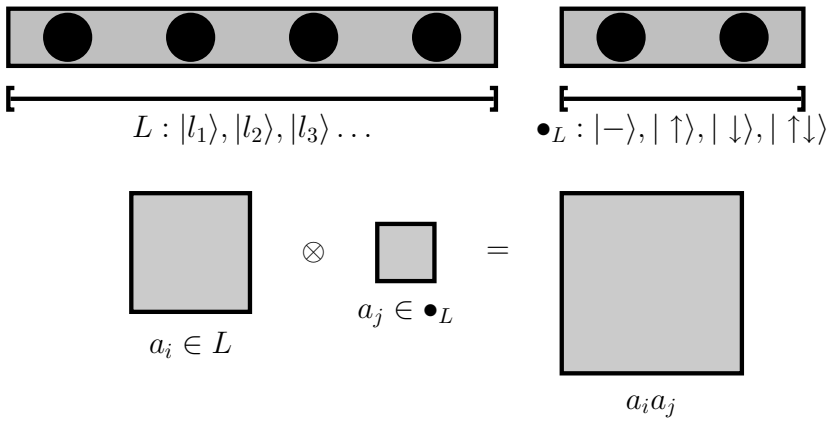


Figure 2.3: Blocking step in DMRG. L block (spanned by $\{l\}$ of size M) is blocked with $\bullet_L$ (spanned by $\{\bullet_L\}$ of size 4). Thus the operators are formed through direct products of operators in each block and is represented by matrices of dimension $4M \times 4M$.

If the block L contains M states, after the blocking step the superblock contains $4M$ states $\{l'\}$, as the $\bullet_l$ contains the complete Fock space of size 4 ($|-\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle$) of the dot. The Hamiltonian of the superblock $L + \bullet_L$

is given by $H_{L'}$

$$\begin{aligned} H_{L'} &= H_L + H_{\bullet_L} + \sum_{i\in L, j\in \bullet_L} t_{ij}(a_i^\dagger a_j + a_j^\dagger a_i) \\ &+ \sum_{ijk\in L, l\in \bullet_L} (v_{ijkl} - v_{ijlk}) a_i^\dagger a_j^\dagger a_k a_l \\ &+ \sum_{i\in L, jkl\in \bullet_L} (v_{ijkl} - v_{jikl}) a_i^\dagger a_j^\dagger a_k a_l \\ &+ \sum_{ij\in L, kl\in \bullet_L} v_{ijkl}(a_i^\dagger a_j^\dagger a_k a_l + a_k^\dagger a_l^\dagger a_i a_j) \\ &+ \sum_{ik\in L, jl\in \bullet_L} (v_{ijkl} - v_{jikl} - v_{ijlk} + v_{jilk}) a_i^\dagger a_j^\dagger a_k a_l, \end{aligned} \tag{2.2}$$

where the first two terms are the Hamiltonians of the individual block and dot and the third term onwards are the interactions between the block and the dot adjacent to it. The operators in the superblock are made by the direct product between the operators in the block and that of the dot.

2. Solving

The decimation criterion explained earlier requires the creation of the approximate Hamiltonian of the complete system given by $H_{L'R'}$,

$$\begin{aligned} H_{L'R'} &= H_L \otimes H_{\bullet_L} \otimes H_{\bullet_R} \otimes H_R \\ &= Dim(M) * Dim(4) * Dim(4) * Dim(M) \\ &= H_{L'} \otimes H_{R'}. \end{aligned} \tag{2.3}$$

This eigenstate(s) of the approximate Hamiltonian $H_{L'R'}$ are obtained as in Eqns. 2.4 and 2.5. Here, it should be noted that the Hamiltonian is of dimension $16M^2$ as each L' and R' are of dimension $4M$ (see Eqn. 2.3).

$$H_{L'R'}\Psi = \epsilon\Psi \tag{2.4}$$

$$\Psi = c_{l'r'}|l'r'\rangle. \tag{2.5}$$

3. Decimation

In the next step, we decimate or transform or renormalize the superblock subspace of $4M$ into a basis of dimension M. This is the step where the degrees of freedom are reduced while approximately preserving the eigenvalue spectrum associated with the wavefunctions obtained in Eqn. 2.4.

From the approximate wavefunction given by the Eqn. 2.5, the density matrix can be constructed (dropping the $\prime$ notation in Eqn. 2.5 and taking l, r to refer to states in the superblocks $L + \bullet_L$, $R + \bullet_R$)

$$\begin{aligned} D_{ll'} &= \sum_r c_{lr} c_{l'r} \\ &= \sum_i w_i |\theta_i\rangle\langle\theta_i|. \end{aligned} \tag{2.6}$$

The eigenfunctions θ_i corresponding to the lowest M eigenvalues of the density matrix are chosen as the decimated states (indicated by a bar) thus giving the approximate renormalized wavefunction,

$$|\Psi\rangle \approx |\bar{\Psi}\rangle = \sum_{\bar{l}\bar{r}} c_{\bar{l}\bar{r}} |\bar{l}|\bar{r}\rangle. \tag{2.7}$$

The eigenfunction basis of the density matrix forms a fast converging basis for the wavefunction. This can be verified if we define our problem as one of finding the approximate wavefunction and minimizing the error, given by

$$\begin{aligned} S &= ||\Psi\rangle - |\bar{\Psi}\rangle|^2 \\ &= \sum_{lr} (c_{lr} - \sum_i d_i u_l^i v_r^i)^2 \end{aligned} \tag{2.8}$$

where i is of dimension M and l and r are of dimension $4M$. Thus, we are trying to minimize S by optimizing d_i, u_l^i and v_r^i and the solution is produced by the singular value decomposition of $\mathbf{c}$ (with elements c_{lr}),

taking the M largest singular values

$$\mathbf{c} = \mathbf{UDV}^T \tag{2.9}$$

where $\mathbf{U}$ is $4M \times M$, $\mathbf{D}$ is $M \times M$, V is $M \times 4M$. U is row-orthogonal, V is column-orthogonal and $\mathbf{D}$ is a diagonal matrix containing the singular values of $\mathbf{c}$.

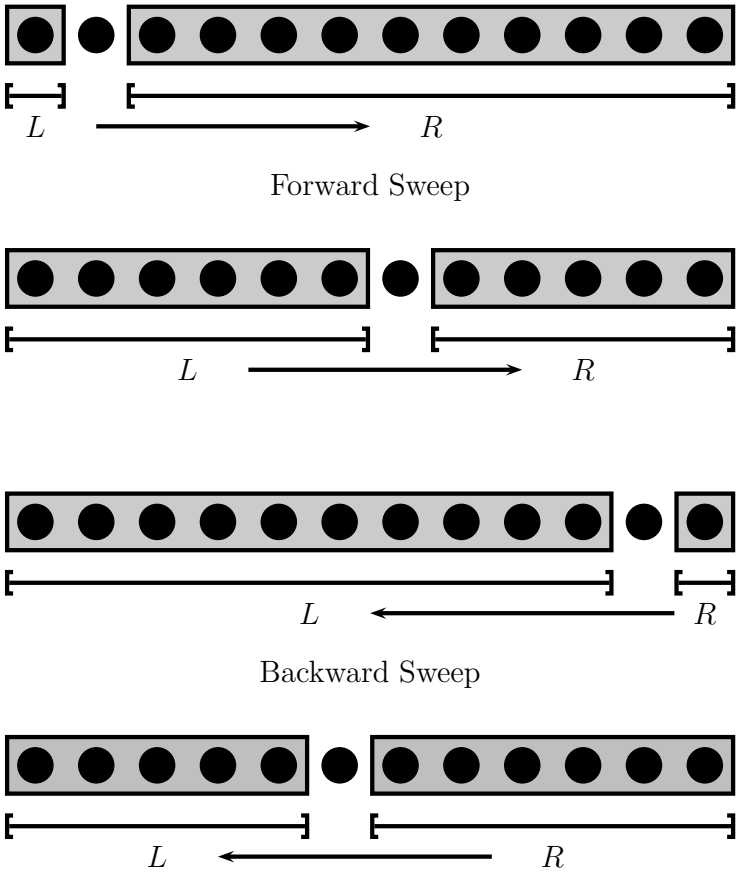


Figure 2.4: The DMRG sweep algorithm consisting of the forward and backward sweep. In the forward sweep L is the system and R is the environment and the system block L grows in size. In the backward sweep the roles of L and R are reversed.

Successive renormalization transformations along the *sweep* are used. We start with a very small left block or system and keep adding sites next to it, successively blocking and decimating as we go over all the sites. This is referred to as a sweep. When we reach one end of the sweep, we reverse and start with a small environment and keep adding sites to it, blocking and decimating. Thus, we have what is called a forward and a backward sweep.

Now, that we know the basic method of applying DMRG to the quantum chemical system, there are some specific problems that need to be addressed for quantum chemical systems due to their special complicated structure.

2.4.1 Dealing with complicated Hamiltonians

From the expression for the quantum chemical Hamiltonian, Eqn. 2.1 it is clear that the Hamiltonian contains $O(k^4)$ terms (k being the number of basis functions or site functions) because of the two electron part $v_{ijkl}a_i^\dagger a_j^\dagger a_k a_l$. Thus, at each step of the sweep, there is need to transform and store $O(k^4)$ such operators each of which has a $M \times M$ matrix representation. This would be very expensive computationally as well as for memory and storage costs if this is done in a brute force way.

Note that we cannot simply store the summation $\sum_{ijkl} v_{ijkl}a_i^\dagger a_j^\dagger a_k a_l$ over all i, j, k, l because we need to update the representation of the operators a_i, a_j etc. during the sweep.

However, there is a common procedure of summing over the terms in the Hamiltonian of a block to create computational intermediates, that is used to

reduce the cost of the quantum chemistry DMRG algorithm. Since the lattice sites are divided into sites on the left and right blocks, the computational intermediates can be classified broadly into three types:

- All four indices in one block, i.e. the operator is located on any one of the component blocks L or R and can be stored as a single floating point matrix by summing over the indices, e.g. $\sum_{ijkl \in L} v_{ijkl} a_i^\dagger a_j^\dagger a_k a_l$.
- 3 indices in one block, e.g. $R_i = \sum_{jkl \in L} w_{ijkl} a_j^\dagger a_k a_l$ which is reduced to $O(k)$ terms on contracting over j, k, l.
- 2 indices in one block, e.g. $P_{ij} = \sum_{kl \in L} v_{ijkl} a_k a_l$ or $Q_{ij} = \sum_{kl \in L} v_{ijkl} a_k^\dagger a_l$, which is reduced to $O(k^2)$ terms.

Using these *complementary operators*, the storage is reduced to $O(k^2 M^2)$ since each of these operators have a $M \times M$ matrix representation. The computational cost of making these complementary operators is $O(M^2 k^3)$ and this is the predominant cost in the blocking step.

Another computationally expensive step arises after blocking, when we have the Hamiltonian $H_{L'R'}$ and we have to construct the Hamiltonian vector product $H_{L'R'} c_{L'R'}$, where c is $16M^2$ in dimension. However, again we can take advantage of the complementary operators and form the Hamiltonian by multiplying these operators, e.g. multiplications like $\sum_{ij} P_{ij} a_i^\dagger a_j^\dagger$ which has a computational cost of $O(M^3 k^2)$ adding up to a total cost of $O(M^3 k^3)$ per DMRG sweep.

To summarize, the most expensive steps in the DMRG algorithm are blocking, Davidson diagonalization and decimation which have a computational scaling of $O(M^2 k^3)$, $O(M^3 k^2)$ and $O(M^3 k^2)$ respectively. Since along each of the sweeps there are $O(k)$ renormalization steps, the total computational cost of the DMRG

algorithm is $O(M^2k^4 + M^3k^3)$ per sweep. Taking into account the optimal combination of complementary operators that must be formed, the memory usage per sweep is $O(M^2k^2)$.

2.4.2 Orbital ordering

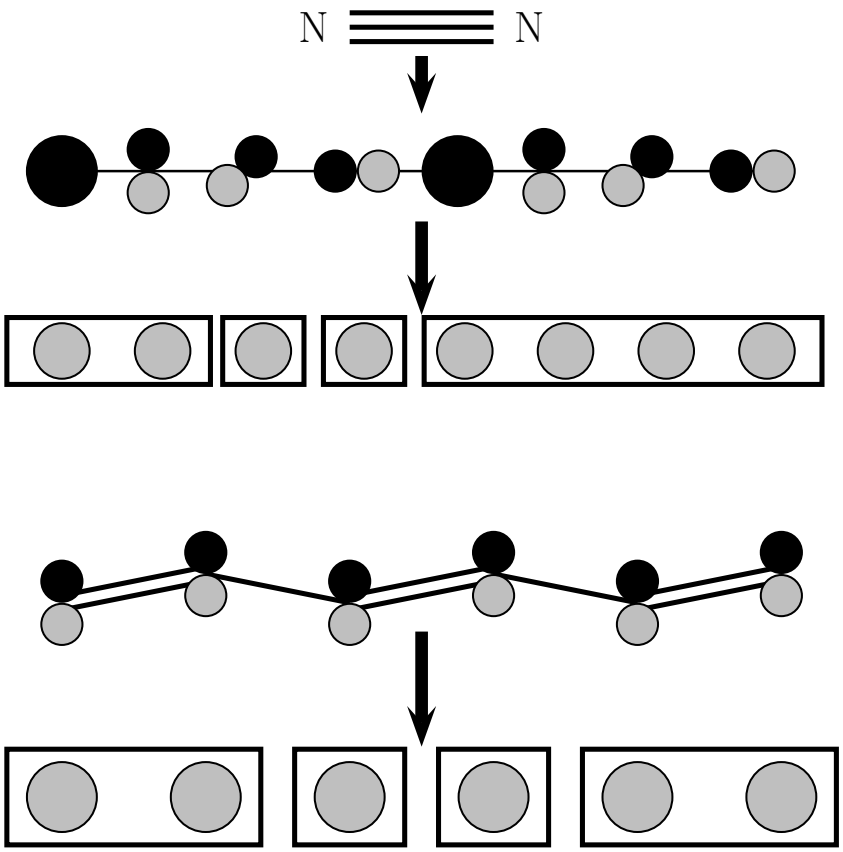


Figure 2.5: Example of orbital ordering in N_2 and C_6H_8. The active space of N_2 consists of $2s$ and $2p$ orbitals. The active space of C_6H_8 consists of the six p_z orbitals each located on a carbon atom.

To start the DMRG sweep, we have to first arrange the orbitals in an one dimen-

sional lattice so that we can start blocking and decimating from one end of the lattice to the other.

This is not always trivial for most quantum chemical systems and especially for complicated active spaces. The orbital ordering is considerably easier if the molecule is pseudo one dimensional. It is further simplified if the active space consists of one orbital per atom, e.g. conjugated polyenes where the active space is generally the π space consisting of the approximately the p_z orbitals on each atom.

When we have to consider more than one dimensional molecules, we have to order the active orbitals in such a way that the nearby lattice sites (corresponding to the orbitals) interact more and the interaction between the lattice sites separated by a longer distance are generally much lower. In this situation, larger M states should be retained to solve for the wavefunction with sufficient accuracy.

There is one more requirement for the successful application of the DMRG ansatz. The orbitals or the site functions should be localized in order to minimize the interaction between distant sites. Different localization schemes are explained in details in Chapter III.

2.4.3 Convergence of the DMRG sweeps

There are two issues of convergence to consider in relation to the DMRG. The first is the fundamental convergence of the truncation procedure as one increases the number of renormalized states used, M. It can be shown that the

error in DMRG energy is proportional to the weight of the discarded states [12]

$$\text{Error} = |E(M) - E_0| \approx \text{const} \sum_i w_i + C. \tag{2.10}$$

The convergence of energy with M can also be shown to be approximately

$$ln|\delta E| \approx -\kappa(ln(M))^2. \tag{2.11}$$

The second issue is the convergence of the iterative sweeps. These generally converge well, but during the decimation step, some random noise sometimes needs to be added to the density matrix so that each state is represented with non-zero weight. As long as this noise is gradually decreased, it does not affect the converged DMRG energy. However, if no noise is added the DMRG solution can get stuck in local minima and does not converge to the correct energy.

2.5 DMRG wavefunction as matrix product states

The DMRG can be understood from a different viewpoint to the one we have so far used, based on the structure of the underlying class of variational wavefunctions, which are sometimes known as Matrix Product States (MPS) [16–18].

The DMRG wavefunction when expressed explicitly in the complete basis expansion of Slater determinants $\{|n_1 n_2 n_3 \dots n_k\rangle\}$ can be written as

$$|\Psi_{DMRG}\rangle = \sum_{\substack{n_1 n_2 n_3 \dots n_k \\ i_1 i_2 i_3 \dots i_{k-1}}} \psi_{i_1}^{n_1} \psi_{i_1 i_2}^{n_2} \psi_{i_2 i_3}^{n_3} \dots \psi_{i_{k-1}}^{n_k} |n_1 n_2 n_3 \dots n_k\rangle \tag{2.12}$$

where $|n_i\rangle$ spans over $|-\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle$.

Each of the functions ψ (other than the ones at the ends) has three indices or ψ is a tensor of rank 3 and thus, can be thought of as a vector or series of matrices.

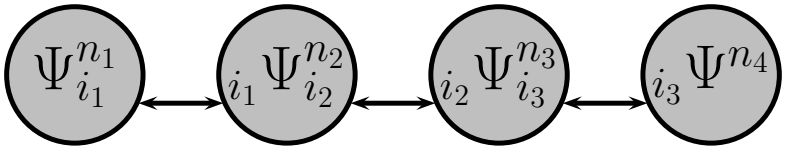


Figure 2.6: Site connectivity in a DMRG site chain. The wavefunction is contracted over the indices i and decimated such that wavefunction is spanned by M determinants.

This way of re-expressing helps us in keeping track of the indices. Also, computationally we are more used to handling vectors and matrices and not higher rank tensors. We can therefore write out the 3-rank tensors as a series of matrix products,

$$|\Psi_{DMRG}\rangle = \sum_{n_1 n_2 n_3 \ldots n_k} [\boldsymbol{\psi^{n_1}}][\boldsymbol{\psi^{n_2}}][\boldsymbol{\psi^{n_3}}]\ldots[\boldsymbol{\psi^{n_k}}] \tag{2.13}$$

and hence the name *Matrix Product States*. In Eqn. 2.13 each of $[\boldsymbol{\psi^{n_j}}]$ is a matrix that has its elements $\psi^{n_j}_{i_{j-1}i_j}$. Here, it should be noted that the matrices that denote the end of the chain of sites, namely $[\boldsymbol{\psi^{n_1}}]$ and $[\boldsymbol{\psi^{n_k}}]$ are of dimension $1 \times M$ and $M \times 1$ respectively and the rest of the matrices in the middle of the chain are of dimension $M \times M$.

In order to understand the analogy with the original DMRG formulation, the above matrices can be easily shown to be formed from a series of singular value decompositions of the FCI tensor $\psi^{n_1 n_2 n_3 \ldots n_k}$ that are related to the singular value decomposition considered in Eqn. 2.9. For example, consider a wavefunction for two sites with FCI coefficients $\psi^{n_1 n_2}$. This wavefunction can be re-expressed

as a singular value decomposition

$$\begin{aligned}\psi_{n_1 n_2} &= \sum_{i_1} R^{n_1}_{i_1} \sigma_{i_1} R^{n_2}_{i_1} \\ &= \sum_{i_1} \psi^{n_1}_{i_1} \psi^{n_2}_{i_1}. \end{aligned} \tag{2.14}$$

Now, if we consider three sites, we can again perform similar successive singular value decompositions

$$\begin{aligned}\psi_{n_1 n_2 n_3} &= \sum_{i_1} R^{n_1}_{i_1} \sigma_{i_1} S^{n_2 n_3}_{i_1} \\ &= \sum_{i_1 i_2} R^{n_1}_{i_1} \sigma_{i_1 i_2} R^{n_2}_{i_1 i_2} R^{n_3}_{i_2} \\ &= \sum_{i_1 i_2} \psi^{n_1}_{i_1} \psi^{n_2}_{i_1 i_2} \psi^{n_3}_{i_2}. \end{aligned} \tag{2.15}$$

Thus, when we consider the complete chain of site functions, we end up with $k-1$ singular value decompositions to get a matrix product state that can be re-written in the form Eqn. 2.12.

2.6 Further properties of the DMRG

We finish by mentioning some further important properties of the DMRG that are relevant to quantum chemistry.

1. Variational nature

   When we start a DMRG sweep, we start with the complete Hilbert space of a small subsystem and as we go along the sweep, we keep adding the Hilbert space of the next site and then decimating or choosing a small

subset of this Hilbert space. Thus, at all steps the DMRG wavefunction is spanned by a subset of the Hilbert space. And therefore, it is variational in nature. With the increase in M or the number of states that are kept after decimation, the DMRG energy reaches the FCI energy variationally.

2. Size-consistency

   The DMRG wavefunction is size-consistent in the localized basis. $|\Psi^{AB}\rangle$ (the wavefunction of the composite system) can be shown to be equal to the product of the wavefunction of the subsystems $|\Psi^{A}\rangle$ and $|\Psi^{B}\rangle$,

$$\begin{aligned}
|\Psi^{AB}\rangle &= \sum_{\substack{n_1 \ldots n_{a+b} \\ i_1 \ldots i_{a+b-1}}} \psi_{i_1}^{n_1} \ldots \psi_{i_{a-1} i_a}^{n_a} \psi_{i_a i_{a+1}}^{n_{a+1}} \ldots \psi_{i_{a+b-1}}^{a_{a+b}} |n_1 \ldots n_a n_{a+1} \ldots n_{a+b}\rangle \\
&= \sum_{\substack{n_1 \ldots n_a \\ i_1 \ldots i_{a-1}}} \psi_{i_1}^{n_1} \ldots \psi_{i_{a-1}}^{n_a} |n_1 \ldots n_a\rangle \sum_{\substack{n_{a+1} \ldots n_{a+b} \\ i_{a+1} \ldots i_{a+b-1}}} \psi_{i_{a+1}}^{n_{a+1}} \ldots \psi_{i_{a+b-1}}^{n_{a+b}} |n_{a+1} \ldots n_{a+b}\rangle \\
&= |\Psi^{A}\rangle |\Psi^{B}\rangle. \qquad (2.16)
\end{aligned}$$

3. Multireference wavefunction for one dimensional systems

   The DMRG wavefunction is multireference as the occupied and virtual orbitals are treated equally. And it is naturally suitable for the treatment of one dimensional or pseudo one dimensional systems as the sites can be naturally ordered along the backbone of the system and are blocked and decimated along one dimension. This gives a physically correct way of blocking the orbitals that are spatially close to each other and therefore, has the largest correlation. Using localized orbitals, the number of states retained can be virtually kept constant with the length of the molecule (if it is pseudo one dimensional) without any loss of accuracy.

BIBLIOGRAPHY

[1] L. P. Kadanoff, Real space renormalization techniques, in *Statistical Physics: Statics Dynamics and Renormalization*, page 291, World Scientific, Singapore, 2000.

[2] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[3] K. G. Wilson, Rev. Mod. Phys. **55**, 583 (1983).

[4] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[5] R. M. Noack and S. R. White, Phys. Rev. Lett. **68**, 3487 (1992).

[6] R. Noack and S. White, The density matrix renormalization group, in *The Density Matrix Renormalization Group : A New Numerical Method in Physics*, Springer, 1999.

[7] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).

[8] K. Hallberg, Density matrix renormalization: A review of the method and its applications, in *Theoretical Methods for Strongly Correlated Electrons*, edited by D. Sénéchal, A.-M. Tremblay, and C. Bourbonnais, CRM Series in Mathematical Physics, Springer, New York, 2003.

[9] G. Fano, F. Ortolani, and L. Ziosi, J. Chem. Phys. **108**, 9246 (1998).

[10] S. R. White and R. L. Martin, J. Chem. Phys. **110**, 4127 (1999).

[11] A. O. Mitrushenkov, G. Fano, F. Ortolani, R. Linguerri, and P. Palmieri, J. Chem. Phys. **115**, 6815 (2001).

[12] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).

[13] G. K.-L. Chan, J. Chem. Phys. **120**, 3172 (2004).

[14] Ö. Legeza, J. Röder, and B. A. Hess, Mol. Phys. **101**, 2019 (2003).

[15] G. Moritz, B. A. Hess, and M. Reiher, J. Chem. Phys. **122**, 024107 (2005).

[16] S. Rommer and S. Östlund, Thermodynamic limit and matrix-product states, in *The Density Matrix Renormalization Group : A New Numerical Method in Physics*, Springer, 1999.

[17] F. Verstraete, D. Porras, and J. I. Cirac, Phys. Rev. Lett. **93**, 227204 (2004).

[18] G. K.-L. Chan et al., Progress in Theoretical Chemistry and Physics **18**, 49 (2008).

CHAPTER 3

DENSITY MATRIX RENORMALIZATION GROUP SELF-CONSISTENT FIELD

(Some parts of this chapter are taken from Ref. [1].)

3.1 Introduction

The basic idea of the DMRG wavefunction was explained in Chapter II. Since DMRG gives us a variational ansatz to solve for the eigenstates of the quantum chemical Hamiltonian with arbitrary accuracy, we can use it as an alternative for full CI calculations. However, in quantum chemistry the full CI or DMRG calculations in our case is done only in the active space in order to reduce the computational cost. That is we start by dividing the orbitals into core, active and virtual orbitals as defined in Chapter I. The active space consists of the valence orbitals where the most important chemistry occurs and therefore accounts for the majority of the chemically relevant correlations. DMRG calculations are carried out with the orbitals in the active space.

In most applications of the DMRG to quantum chemistry so far [2–5], the active space of interest has been easy to identify, i.e. there is a good core-valence and valence-Rydberg separation, either for energetic or symmetry reasons, allowing the DMRG to be used with such an active space as a direct substitute for complete active space configuration interaction (CASCI). In general however, we cannot always identify the active orbitals in a simple way, and thus there is a need for an *orbital optimized* DMRG, where the active space is determined self-consistently by energy minimization, in much the same way as in the complete

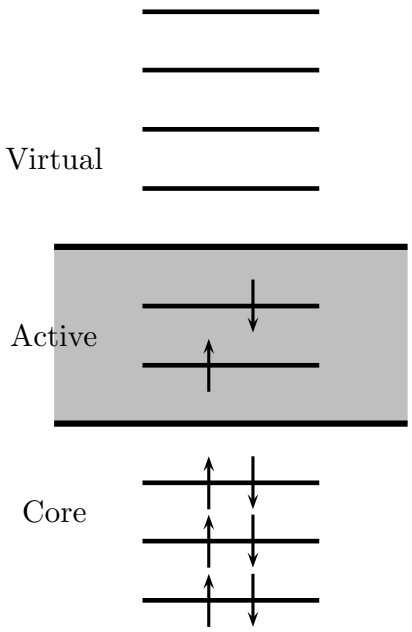


Figure 3.1: The orbitals in a system are divided into core, active and virtual spaces. The core consists of the low energy orbitals that are predominantly doubly occupied. The active space consists of the valence orbitals that have a variable occupancy and are the most important in chemical reactions. The virtual orbitals are very high energy orbitals that are almost never occupied (see Chapter I).

active space self-consistent field (CASSCF) method [6, 7]. In this method we start with a reasonable guess of the active space by looking at the energy and the symmetry of the orbitals and self-consistently solve for the optimal active space. The resulting orbital optimized DMRG will be referred to as the **DMRG-CASSCF** method.

Orbital optimization is a straightforward idea however, its efficient and practical implementation has to be tailored to the underlying many body wavefunc-

tion ansatz and the details of our implementation are explained in this chapter. Our implementation ties together calculating the DMRG wavefunction in the active space, obtaining the reduced density matrices and using these information to rotate the orbitals in order to minimize the energy. Since DMRG is a local ansatz, we also have to consider the various localization schemes. We have a fully parallelized implementation in order to enable us to do large scale calculations.

3.2 Overview of orbital optimization

Starting from the electronic Hamiltonian, specified by the one- and two-electron integrals t_{ij} and v_{ijkl}

$$H = \sum_{ij} t_{ij} a_i^\dagger a_j + \sum_{ijkl} v_{ijkl} a_i^\dagger a_j^\dagger a_k a_l \tag{3.1}$$

an *ab-initio* quantum chemical method provides a wavefunction Ψ that approximates a target eigenstate of H. From Ψ we define the one- and two-particle density matrix elements $\gamma_{ij}, \gamma_{ijkl}$

$$\begin{aligned} \gamma_{ij} &= \langle \Psi | a_i^\dagger a_j | \Psi \rangle \\ \gamma_{ijkl} &= \langle \Psi | a_i^\dagger a_j^\dagger a_k a_l | \Psi \rangle \end{aligned} \tag{3.2}$$

and the energy expectation value $\langle \Psi | H | \Psi \rangle$ can be written as

$$E = \sum_{ij} t_{ij} \gamma_{ij} + \sum_{ijkl} v_{ijkl} \gamma_{ijkl} \tag{3.3}$$

Orbital rotation corresponds to a unitary transformation of the wavefunction effected by an operator e^A, where A has the single-particle excitations

$$A = \sum_{ij} A_{ij} a_i^\dagger a_j \tag{3.4}$$

and $A_{ij} = -A^*_{ji}$. After orbital rotation, the transformed wavefunction $\bar{\Psi}$ and energy $\bar{E}$ are

$$\begin{aligned} \bar{\Psi} &= e^A \Psi \\ \bar{E} &= \langle \Psi e^{-A} | H | e^A \Psi \rangle \end{aligned} \tag{3.5}$$

But one can also consider the unitary operator to act on the Hamiltonian rather than the wavefunction, and from this equivalent point of view, we have a transformed $\bar{H}$ and energy expressions

$$\bar{H} = e^{-A} H e^A, \tag{3.6}$$

$$\bar{E} = \langle \Psi | \bar{H} | \Psi \rangle. \tag{3.7}$$

The transformed Hamiltonian $\bar{H}$ has the same form as the original Hamiltonian (Eqn. 3.1) but with modified integrals $\bar{t}_{ij}$ and $\bar{v}_{ijkl}$ that reflect the rotated orbitals

$$\begin{aligned} \bar{t}_{ij} &= \sum_{i'j'} U^*_{ii'} U_{jj'} t_{i'j'} \\ \bar{v}_{ijkl} &= \sum_{i'j'k'l'} U^*_{ii'} U^*_{jj'} U_{kk'} U_{ll'} v_{i'j'k'l'} \end{aligned} \tag{3.8}$$

where U is the coefficient matrix e^A. Thus we can rewrite the energy after orbital rotation in terms of the original one- and two-particle density matrices and the modified integrals

$$\bar{E} = \sum_{ij} \bar{t}_{ij} \gamma_{ij} + \sum_{ijkl} \bar{v}_{ijkl} \gamma_{ijkl}. \tag{3.9}$$

Thus, the basic procedure of optimizing the orbital in an *ab-initio* wavefunction is as follows:

1. From the *ab-initio* method obtain Ψ corresponding to the given H and form the density matrices $\gamma_{ij}, \gamma_{ijkl}$.

2. Determine an orbital rotation step e^A, and form the new Hamiltonian $\bar{H} = e^{-A} H e^A$ from the transformed integrals.

3. Goto 1. and loop until convergence in Ψ.

Note that in the above, the orbital degrees of freedom and the other ansatz degrees of freedom in Ψ are alternately optimized in steps (1), (2). More sophisticated approaches which couple orbital rotations with changes in the other ansatz degrees of freedom can be envisaged (as are employed in multi-configurational self-consistent field methods [8–10]), but we have used the above simple strategy to optimize the orbitals in the DMRG wavefunction. The conceptual task is then twofold:

- How do we calculate the one- and two-particle density matrices from the DMRG wavefunction?
- What method should we use to select our orbital rotation steps and to construct the transformed Hamiltonian?

3.3 One- and two-site DMRG algorithms

Before we go on any further, we will digress from the discussion and define more clearly the different kinds of DMRG algorithm based on different blocking schemes. The algorithm described in Chapter II (Sec. 2.4) is a "two-site" DMRG algorithm, while the wavefunction written in Sec. 2.5 corresponds to a "one-site" algorithm. Here we present both the wavefunction and algorithm expressions for the one- and two-site DMRG side by side as they will be crucial to

understanding the one- and two-body density matrix calculation as described later.

3.3.1 Two-site algorithm

In the two-site algorithm the DMRG wavefunction can be written as

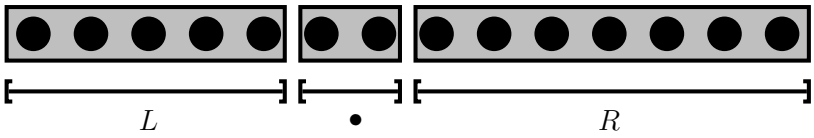
$$|\Psi\rangle = \sum_{n_1 n_2 \ldots n_p} \boldsymbol{L}^{n_1}\boldsymbol{L}^{n_2}\ldots\boldsymbol{L}^{n_p}|n_1 n_2 \ldots n_p\rangle \otimes |n_{p+1}\rangle \otimes |n_{p+2}\rangle \otimes \sum_{n_{p+3} n_{p+4} \ldots n_k} \boldsymbol{C}^{n_p n_{p+1} n_{p+2} n_{p+3}} \boldsymbol{R}^{n_{p+3}} \boldsymbol{R}^{n_{p+4}} \ldots \boldsymbol{R}^{n_k} |n_{p+3} n_{p+4} \ldots n_k\rangle, \quad (3.10)$$

where we have defined all the site functions or renormalization matrices on the left block as $\boldsymbol{L}^{n_i}$, those on the right block as $\boldsymbol{R}^{n_i}$ and $\boldsymbol{C}^{n_p n_{p+1} n_{p+2} n_{p+3}}$ is the correlation between the blocks and the sites in between (explained in more details in the next section). Here, we see that up to p sites are on the left block, $p+1$ and $p+2$ form the left and right sites and the $p+3$ to k sites are in the right block. Thus, at each step the wavefunction is defined in a space of dimension $M \times 4 \times 4 \times M$ as the left and right blocks have M states in them and the sites in between have the complete Fock space (4 states each). This algorithm was the first to be developed and exhibits robust convergence.

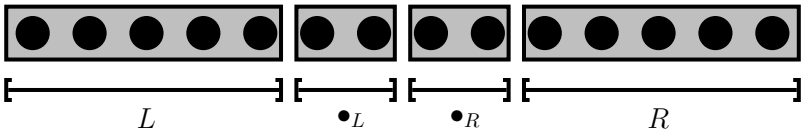
3.3.2 One-site algorithm

The DMRG wavefunction as defined by the one-site algorithm is

$$|\Psi\rangle = \sum_{n_1 n_2 \ldots n_p} \boldsymbol{L}^{n_1}\boldsymbol{L}^{n_2}\ldots\boldsymbol{L}^{n_p}|n_1 n_2 \ldots n_p\rangle \otimes |n_{p+1}\rangle \otimes \sum_{n_{p+2} n_{p+3} \ldots n_k} \boldsymbol{C}^{n_p n_{p+1} n_{p+2}} \boldsymbol{R}^{n_{p+2}} \boldsymbol{R}^{n_{p+3}} \ldots \boldsymbol{R}^{n_k} |n_{p+2} n_{p+3} \ldots n_k\rangle, \quad (3.11)$$



(a) One site DMRG block



(b) Two site DMRG block

Figure 3.2: The one- and two-site algorithms in DMRG. They are different in the number of sites (dots) in between the L and R blocks. The lattice sites in the figure correspond to spin orbitals, i.e. a pair of these spin orbitals make a spatial orbital or dot.

where we have defined all the site functions or renormalization matrices on the left block as $\boldsymbol{L}^{n_i}$, those on the right block as $\boldsymbol{R}^{n_i}$ and $\boldsymbol{C}^{n_p n_{p+1} n_{p+2}}$ is the correlation between the blocks and the site in between (explained in more details in the next section). Here, we see that up to p sites are on the left block, $p+1$ form the dot and $p+2$ to k sites are in the right block.

From Eqn. 3.11 it follows that the one-site DMRG wavefunction is defined in a space of dimension $M \times 4 \times M$ as the left and right blocks have M states in them and the site in between has a complete Fock space of 4. The dimension of the Hamiltonian in the Davidson diagonalization is $4M^2 \times 4M^2$ which is 16 times smaller than in the two-site algorithm. The dimensions of the density matrix ($4M \times M$), ensure that there can only be M non-zero eigenvalues and thus, the choice of eigenvectors are trivial. Consequently, there are no trunca-

tion errors. Thus, the wavefunction can be exactly expressed in either of the forms $\boldsymbol{L}^{n_1}\boldsymbol{L}^{n_2}\ldots\boldsymbol{L}^{n_p}|n_1 n_2\ldots n_p\rangle \otimes \boldsymbol{C}^{n_p n_{p+1} n_{p+2}}|n_{p+1}\rangle \otimes \boldsymbol{R}^{n_{p+2}}\ldots\boldsymbol{R}^{n_k}|n_{p+2}\ldots n_k\rangle$ and $\boldsymbol{L}^{n_1}\boldsymbol{L}^{n_2}\ldots\boldsymbol{L}^{n_{p-1}}|n_1 n_2\ldots n_{p-1}\rangle \otimes \boldsymbol{C}^{n_{p-1} n_p n_{p+1}}|n_p\rangle \otimes \boldsymbol{R}^{n_{p+1}}\ldots\boldsymbol{R}^{n_k}|n_{p+1}\ldots n_k\rangle$ corresponding to successive block configurations in the sweep. In the one-site algorithm, the energy converges monotonically within a sweep in contrast to the two-site algorithm where the minimum energy is at the middle of the sweep. Consequently at convergence the DMRG wavefunction is the same at any block configuration and this is crucial to correctly evaluate the one- and two-body reduced density matrices in the procedure we have described later. However, in spite of all these very attractive properties of the one-site algorithm, it is very prone to converging to a local minima. In our DMRG-CASSCF procedure, we have used the one-site algorithm but, have used rather large numerical noise to ensure that the wavefunction does not get trapped in a local minima and reaches the correct energy.

3.4 Canonical structure of the DMRG wavefunction

To make the properties of the one-site DMRG algorithm introduced in the last section more explicit (in particular the relationship between the algorithm and wavefunction formulation of the DMRG), we now go into greater detail to describe the relationship between DMRG wavefunctions at different block configurations in a sweep.

The DMRG ansatz is invariant to transformations of the site functions of the form ($\boldsymbol{\psi}^{n_p} \rightarrow \boldsymbol{\psi}^{n_p}\mathbf{U}$, $\boldsymbol{\psi}^{n_{p+1}} \rightarrow \mathbf{U}^{\dagger}\boldsymbol{\psi}^{n_{p+1}}$) and thus it is useful to define a *canonical* form of the DMRG wavefunction that eliminates this freedom. In practice, this

canonical representation is used in all DMRG calculations, and it is also the representation in which the link between the DMRG wavefunction and the traditional renormalization group language is most direct. In essence, the canonical form of the wavefunction at a given site corresponds to the familiar expression for the DMRG wavefunction where it is expanded in the product basis of the left and right blocks separated by the site [11, 12].

To obtain the canonical form, we choose a specific site, say p, around which to canonicalize. Then the site p canonical form is given as

$$|\Psi\rangle = \sum_{n_1 \ldots n_p \ldots n_k} \boldsymbol{L}^{n_1} \ldots \boldsymbol{L}^{n_{p-1}} \boldsymbol{C}^{n_p} \boldsymbol{R}^{n_{p+1}} \ldots \boldsymbol{R}^{n_k} |n_1 \ldots n_p \ldots n_k\rangle. \tag{3.12}$$

We label the site functions to the left of p by L and those to the right by R. The degeneracy (invariance to transformation) of the original ansatz (Eqn. 3.12) mentioned above is lifted by requiring the L and R site functions to be orthogonal projection matrices in the following sense

$$\sum_{l n_q} L^{n_q}_{l l'} L^{n_q}_{l l''} = \delta_{l' l''} \tag{3.13}$$

$$\sum_{r n_q} R^{n_q}_{r' r} R^{n_q}_{r'' r} = \delta_{r' r''} \tag{3.14}$$

i.e. by grouping together the ln_q indices to form the row index of a $4M \times M$ matrix, each L site function is orthogonal with respect to its M columns, while by grouping together the rn_q indices to form the column index of a $M \times 4M$ matrix, each R site function is orthogonal with respect to its M rows.

The link between the canonical form and the original RG formulation appears when we combine the L site functions $\boldsymbol{L}^{n_1} \ldots \boldsymbol{L}^{n_{p-1}}$ with the basis states $|n_1 \ldots n_{p-1}\rangle$, and the R site functions $\boldsymbol{R}^{n_{p+1}} \ldots \boldsymbol{R}^{n_k}$ with the basis states $|n_{p+1} \ldots n_k\rangle$, to define renormalized left and right many body spaces $\{l_{p-1}\}$,

$\{r_{p+1}\}$

$$|l_{p-1}\rangle = \sum_{\substack{n_1 \ldots n_{p-1} \\ l_1 \ldots l_{p-2}}} L^{n_1}_{l_1} \ldots L^{n_{p-1}}_{l_{p-2} l_{p-1}} |n_1 \ldots n_{p-1}\rangle, \tag{3.15}$$

$$|r_{p+1}\rangle = \sum_{\substack{n_{p+1} \ldots n_k \\ r_{p+2} \ldots r_k}} R^{n_{p+1}}_{r_{p+1} r_{p+2}} \ldots R^{n_k}_{r_k} |n_{p+1} \ldots n_k\rangle. \tag{3.16}$$

Since the dimension of the left basis in Eqn. 3.15 is M (i.e. the dimension of the auxiliary index l_{p-1}) and similarly for the right basis, the site functions $\boldsymbol{L}^{n_1} \ldots \boldsymbol{L}^{n_{p-1}}$ and $\boldsymbol{R}^{n_{p+1}} \ldots \boldsymbol{R}^{n_k}$ define a projective transformation or *renormalization* from the many body spaces $\{n_1\} \otimes \ldots \otimes \{n_{p-1}\}$ and $\{n_{p+1}\} \otimes \ldots \otimes \{n_k\}$ to the left and right spaces, $\{l_{p-1}\}$, $\{r_{p+1}\}$, respectively. Then in the renormalized representation, $C^{n_p}_{l_{p-1} r_p}$ gives the coefficients of expansion of the wavefunction $|\Psi\rangle$, i.e.

$$|\Psi\rangle = \sum_{l_{p-1} n_p r_p} C^{n_p}_{l_{p-1} r_p} |l_{p-1} n_p r_p\rangle. \tag{3.17}$$

This is just the RG expression for the one-site DMRG wavefunction, in the product space of a renormalized left "block", a site p, and a renormalized right "block". Thus in the usual DMRG language, the site p canonical form corresponds to the DMRG wavefunction in the basis associated with the block configuration $\boxed{\bullet_1 \ldots \bullet_{p-1}}\ \bullet_p\ \boxed{\bullet_{p+1} \ldots \bullet_k}$.

A one-site DMRG wavefunction expressed in the canonical form of a given site p can always be expressed in the canonical form for any other site (or using the traditional DMRG language, the DMRG wavefunction for a given one-site block configuration can always be expressed in the basis of any other one-site block configuration along a sweep). Since we are simply re-expressing the same wavefunction in a different basis, the coefficients C and site-functions L, R at different sites are related. To see the link explicitly, we compare the canonical

forms at adjacent sites p, $p+1$

$$\begin{aligned} |\Psi\rangle &= \sum_{n_1 \ldots n_p \ldots n_k} \boldsymbol{L}^{n_1} \ldots \boldsymbol{L}^{n_{p-1}} \boldsymbol{C}^{n_p} \boldsymbol{R}^{n_{p+1}} \boldsymbol{R}^{n_{p+2}} \ldots \boldsymbol{R}^{n_k} |n_1 \ldots n_p \ldots n_k\rangle \\ &= \sum_{n_1 \ldots n_p \ldots n_k} \boldsymbol{L}^{n_1} \ldots \boldsymbol{L}^{n_{p-1}} \boldsymbol{L}^{n_p} \boldsymbol{C}^{n_{p+1}} \boldsymbol{R}^{n_{p+2}} \ldots \boldsymbol{R}^{n_k} |n_1 \ldots n_p \ldots n_k\rangle, \end{aligned} \tag{3.18}$$

which yields the relation

$$\boldsymbol{L}^{n_p} \boldsymbol{C}^{n_{p+1}} = \boldsymbol{C}^{n_p} \boldsymbol{R}^{n_{p+1}}. \tag{3.19}$$

Now say we are given $\boldsymbol{C}^{n_p} \boldsymbol{R}^{n_{p+1}}$ from the site p canonical form, and we wish to determine $\boldsymbol{L}^{n_p} \boldsymbol{C}^{n_{p+1}}$ for the site $p+1$ canonical form, where $\boldsymbol{L}^{n_p}$ satisfies the orthogonality conditions (Eqn. 3.13). We can obtain such a $\boldsymbol{L}^{n_p}$ solution of Eqn. 3.19 together with $\boldsymbol{C}^{n_{p+1}}$ from the singular value decomposition (SVD) of $\boldsymbol{C}^{n_p}$, viewed as the $4M \times M$ matrix with row indices $l_{p-1} n_p$, column indices r_{p+1} and M singular values σ_{l_p}

$$C^{n_p}_{l_{p-1}, r_{p+1}} = \sum_{l_p} L^{n_p}_{l_{p-1}, l_p} \sigma_{l_p} V_{l_p r_{p+1}} \tag{3.20}$$

$$C^{n_{p+1}}_{l_p, r_{p+2}} = \sum_{r_{p+1}} \sigma_{l_p} V_{l_p r_{p+1}} R^{n_{p+1}}_{r_{p+1}, r_{p+2}}. \tag{3.21}$$

The above transformation between canonical forms at adjacent sites corresponds directly to the transformation between block configurations during the sweep algorithm in the DMRG. In particular, Eqn. 3.20 corresponds to the determination of the basis of the renormalized block $\boxed{\bullet_1 \ldots \bullet_{p+1}}$ from the density matrix eigenvectors of the superblock $\boxed{\bullet_1 \ldots \bullet_p}\ \bullet_{p+1}$, while Eqn. 3.21 corresponds to the wavefunction transformation used to generate the guess at a given block configuration from that at the previous configuration. We see therefore, that an *exact* transformation between canonical forms at different sites is only possible with the *one-site* DMRG ansatz and not in the two-site DMRG discussed in Sec. 3.3.

3.5 Evaluation of one- and two-particle density matrices in DMRG

Given the DMRG wavefunction, the next step is to find the one- and two-particle density matrices. Naively, one would think that the direct evaluation of the one- and two-particle density matrices should require k^2 and k^4 terms respectively and thus, the cost of saving these operators would be $O(M^2k^4)$. This is prohibitively expensive. Here, it should be noted that unlike the Hamiltonian, where we could have contracted over all i, j, k, l indices to get $\sum_{ijkl} v_{ijkl} a_i^\dagger a_j^\dagger a_k a_l$, for the evaluation of two-particle reduced density matrix (2RDM) we actually need *all* of the $a_i^\dagger a_j^\dagger a_k a_l$ terms . Thus, we need to find some computationally cheaper way of calculating the 2RDM from the DMRG wavefunction.

In the algorithm for the evaluation of the two-particle density matrix, we take advantage of the canonical representations at the *different* sites, and certain computations which are difficult in one canonical representation become computationally cheaper in another. When changing between representations, we are simply re-expressing the same wavefunction in a different step of the sweep and exploiting the relationship between the wavefunction coefficient C and the transformation matrices L, R at different steps of the sweep algorithm.

The pseudo-code for the two-particle RDM evaluation is given in Algs. (1), (2). Alg. (1) describes how to partition the evaluation of different density matrix elements amongst the block configurations as we traverse a DMRG sweep.

The different functions that are put together in Alg. (1) to form the complete two-particle density matrix can be understood by noting that there are a few

Algorithm 1: Creation of the two-particle density matrix by assembling across a DMRG sweep.

special treatment for first configuration $\boxed{\bullet_1}\ \bullet_2\ \boxed{\bullet_3 \dots \bullet_k}$

$left$= site 1, $sitep$= site 2, $right$= sites $3 \dots k$

COMPUTE(4, 0, 0, $left$, $sitep$, $right$)

COMPUTE(3, 1, 0, $left$, $sitep$, $right$)

COMPUTE(3, 0, 1, $left$, $sitep$, $right$)

COMPUTE(2, 1, 1, $left$, $sitep$, $right$)

sweep through block configurations $\boxed{\bullet_1 \dots \bullet_{p-1}}\ \bullet_p\ \boxed{\bullet_{p+1} \dots \bullet_k}$

for $sitep$= 2 to k-1 **do**

$left$= sites $1 \dots p-1$, $right$= sites $p+1 \dots k$

COMPUTE(1, 2, 1, $left$, $sitep$, $right$)

COMPUTE(2, 1, 1, $left$, $sitep$, $right$)

COMPUTE(2, 2, 0, $left$, $sitep$, $right$)

COMPUTE(1, 3, 0, $left$, $sitep$, $right$)

COMPUTE(0, 3, 1, $left$, $sitep$, $right$)

COMPUTE(0, 4, 0, $left$, $sitep$, $right$)

end for

special treatment for final configuration $\boxed{\bullet_1 \dots \bullet_{k-2}}\ \bullet_{k-1}\ \boxed{\bullet_k}$

$left$= sites $1 \dots k-2$, $sitep$= site $k-1$, $right$= site k

COMPUTE(0, 0, 4, $left$, $sitep$, $right$)

COMPUTE(0, 1, 3, $left$, $sitep$, $right$)

COMPUTE(1, 0, 3, $left$, $sitep$, $right$)

COMPUTE(0, 2, 2, $left$, $sitep$, $right$)

COMPUTE(2, 0, 2, $left$, $sitep$, $right$)

COMPUTE(1, 1, 2, $left$, $sitep$, $right$)

COMPUTE(1, 2, 1, $left$, $sitep$, $right$)

Algorithm 2: COMPUTE$(nl, np, nr, left, sitep, right)$. Note $nl, np, nr \leq 2$ and $nl + np + nr = 4$, i.e. the number of indices in the two-particle density matrix γ.

for all opl= operators with nl indices on block $left$ **do**

(If parallel, loop only over opl stored on current proc)

for all opp= operators with np indices on block $sitep$ **do**

for all opr= operators with nr indices on block $right$ **do**

$\gamma(np, nl, nr) = parity(opl, opp, opr) \times \langle \Psi | opl \otimes opp \otimes opr | \Psi \rangle$

end for

end for

end for

(If parallel, accumulate contributions from all procs to root processor)

major types of density matrix elements depending on how many operators are situated on different lattice sites.

- γ_{ijkl} such that $i = j = k = l$ which give rise to COMPUTE functions with arguments $[4, 0, 0]$, $[0, 4, 0]$ and $[0, 0, 4]$ (in the first configuration, sweep through the block and final configuration respectively).
- γ_{ijkl} such that $i = j = k \neq l$ which give rise to COMPUTE functions with arguments $[3, 1, 0]$ and $[3, 0, 1]$ in the first configuration, $[1, 3, 0]$ in the sweep through the block, and $[0, 1, 3]$ and $[1, 0, 3]$ in the final configuration.
- γ_{ijkl} such that $i = j \neq k \neq l$ which give rise to COMPUTE functions with arguments $[2, 1, 1]$ in the first configuration, $[1, 2, 1]$ and $[2, 2, 0]$ in the sweep through the block, and $[2, 0, 2]$, $[1, 1, 2]$ and $[1, 2, 1]$ in the final configuration.
- γ_{ijkl} such that $i = j \ \& \ k = l \ \& \ i \neq k$ which give rise to COMPUTE function

with arguments $[0, 2, 2]$ in the sweep through the block.

- γ_{ijkl} such that $i \neq j \neq k \neq l$ which give rise to the most complicated COMPUTE function with arguments $[2, 1, 1]$ in the sweep through the block.

The actual calculation of the density matrix elements is carried out by the function COMPUTE in Alg. (2), which computes all density matrix elements that may be assembled from nl index operators on the left block, np index operators on site p, and nr index operators on the right block.

The resultant computational complexity can be understood if we look more closely at the density matrix element $\gamma_{ijkl}\ \forall\ i \neq j \neq k \neq l$ since these are the most computationally expensive elements. In the pseudo-code it is calculated using the COMPUTE function with $[2, 1, 1]$ arguments, i.e. the left block has 2 indices and the site (dot) and the right block have one index each (This can be thought of as 2 indices per block, viewing the $R + \bullet_R$ as a super R block.). That is we choose a block configuration such that i, j lie in the left block, k in the dot in between the blocks and l in the right block, i.e. $\boxed{\ldots \bullet_i \ldots \bullet_j \ldots}\ \bullet_k\ \boxed{\ldots \bullet_l \ldots}$. The corresponding matrix element may then be evaluated using $a_i^\dagger a_j^\dagger$ on the left block, and a_k on site and a_l on right block, and thus no operator matrices with more than two orbital indices appear on either block (see Fig. 3.3).

By the appropriate choice of partitioning between the left and right blocks, we can arrange things such that we never manipulate operators with more than two orbital labels on either the left or right blocks for any $ijkl$. During a DMRG sweep we iterate through all block configurations where the dividing site $\bullet_p$ ranges from site 2 to site $k-1$. At each block configuration, we then evaluate all the two-particle density matrix elements which do not require more than two-index operators on either the left or right blocks, and assemble the contributions

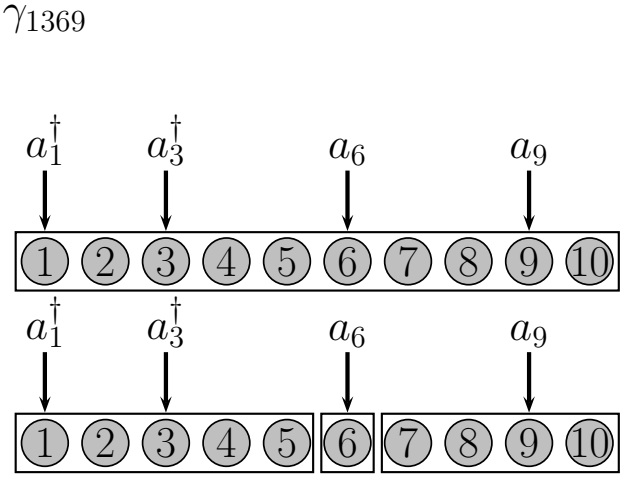


Figure 3.3: Evaluation of a 2-RDM element γ_{1369}. We can obtain this element, e.g. at the block configuration where indices $1, 3$ are on the left block and indices $6, 9$ are on the site and right block (corresponding to calling COMPUTE$(2, 1, 1)$ in Alg. 1).

of all the block configurations at the end of the DMRG sweep. The total memory and storage cost is thus reduced to $O(M^2k^2)$.

We consider the different steps in building these elements.

- The creation of the operators $a_i^\dagger a_j^\dagger$ on left block and $a_k a_l$ on site (in between the blocks) and right block which has k^2 matrix multiplications of size $M \times M$ thus giving rise to the computational cost of M^3k^2.
- The final multiplication of these operators on each block to get the final density matrix element $\gamma_{ijkl} = \sum_{A,B} L_{ij}(A, B)R_{kl}(A, B)$ where there are k^4 indices $ijkl$ and A and B are each of dimension M resulting to the computational cost of M^2k^4.

Therefore, the total computational cost is $O(M^2k^4 + M^3k^2)$. The computational

costs of each of the different objects that build up the complete two-particle reduced density matrix are given in Table 3.1.

Therefore, exploiting the canonical property we can formulate an efficient algorithm to evaluate the two-particle density matrix with a total per-sweep computational cost of $O(M^3k^2 + M^2k^4)$ and a memory cost of $O(M^2k^2)$.

An attractive feature of the quantum chemical DMRG algorithm is the high level of parallelizability, which is described in detail in Ref. [13]. In our implementation, the loops over operators in Alg. (2) are trivially parallelized because of how our operators are divided across processors in the original formulation [13]. For example, the dominant computational cost of the two-particle density matrix evaluation comes from COMPUTE$(2, 1, 1, left, sitep, right)$ in Alg. (1), which costs $O(M^3k^2 + M^2k^4)$ per DMRG sweep. However, in our parallel DMRG implementation, the two index operators opl on the left block, namely $a_i^\dagger a_j$ and $a_i a_j$, are divided across the processors, while the corresponding one index operators opp, opr are replicated on all processors, and thus we can easily parallelize over the first opl loop in Alg. (2). This leads to a final computational cost per sweep of $O((M^3k^2 + M^2k^4)/n_p)$ with a communication cost of $O(k^4 \ln n_p)$, where n_p is the number of processors.

3.6 Orbital step and integral transformation

As described earlier, the DMRG wavefunction is primarily efficient at capturing static correlation and consequently we employ an active space DMRG description of the electronic structure, the purpose of the orbital optimization then being to obtain the best form of the active space. Recall that the active space

Table 3.1: Computational cost of various steps for the evaluation of the two-particle density matrix. k is the no. of orbitals and M is the number of states retained after each decimation step in the DMRG algorithm.

Configuration	Args. to the COMPUTE functions	Total computational cost
First	$[4,0,0]$	$O(M^3)$
	$[3,1,0]$	$O(M^3)$
	$[3,0,1]$	$O(M^3k)$
	$[2,1,1]$	$O(M^3k)$
Sweep	$[1,2,1]$	$O(M^3k^2)$
	$[2,1,1]$	$O(M^3k^2 + M^2k^4)$
	$[2,2,0]$	$O(M^3k^2)$
	$[1,3,0]$	$O(M^3k)$
	$[0,3,1]$	$O(M^3k)$
	$[0,4,0]$	$O(M^3)$
Final	$[0,0,4]$	$O(M^3)$
	$[0,1,3]$	$O(M^3)$
	$[1,0,3]$	$O(M^3k)$
	$[0,2,2]$	$O(M^3)$
	$[2,0,2]$	$O(M^3k^2)$
	$[1,1,2]$	$O(M^3k)$
	$[1,2,1]$	$O(M^3k)$

is defined by partitioning the orbitals into three sets, closed-shell orbitals which remain doubly occupied in all DMRG configurations, active orbitals which form the product active space $\{n_1\} \otimes \ldots \otimes \{n_k\}$ in the DMRG wavefunction expansion Eqn. 2.12, and external orbitals, which remain unoccupied in all DMRG configurations. With this partitioning, the active space DMRG wavefunction is determined with respect to the active space Hamiltonian

$$H^{\text{act}} = E^{\text{closed}} + \sum_{ij} t_{ij}^{\text{act}} a_i^\dagger a_j + \sum_{ijkl} v_{ijkl} a_i^\dagger a_j^\dagger a_k a_l \tag{3.22}$$

where indices i, j are limited to the active orbitals and the modified one-particle integrals t_{ij}^{act} and closed-shell energy are given respectively by

$$E^{\text{closed}} = \sum_{c} t_{cc} + \sum_{cc'} (v_{cc'c'c} - v_{cc'cc'}) \tag{3.23}$$

$$t_{ij}^{\text{act}} = t_{ij} + 2 \sum_{c} (v_{iccj} - v_{icjc}) \tag{3.24}$$

where c, c' denote the closed-shell indices.

Orbital optimization chooses the best form of the active orbitals by minimizing the energy of the DMRG wavefunction with respect to the active and closed-shell orbitals. This is the basic idea behind the complete active space self consistent field (CASSCF) description of electronic structure. In CASSCF, the active space wavefunction is the exact eigenfunction of the active space Hamiltonian Eqn. 3.22 and is thus invariant with respect to active-active orbital rotations. In the corresponding orbital optimized DMRG-CASSCF, the accuracy of our active space DMRG wavefunction depends on the size of M.

The algorithm we use for orbital optimization is an Augmented Hessian Newton Raphson scheme similar to that used in modern CASSCF implementations [8–10]. The orbital rotations are parametrized by the anti-hermitian amplitudes A in Eqn. 3.4, and the derivative with respect to these amplitudes is evaluated

from the one- and two-particle density matrices from the DMRG calculation. However, as the DMRG enables the use of larger active spaces than in traditional CASSCF studies and consequently we can expect to have a larger number of correlating external and closed-shell orbitals, we have focused on an efficient parallel implementation of the orbital optimization. Here the primary task is to parallelize the four-index transformation which is performed after each orbital rotation to generate the two-electron integrals in the basis of the rotated orbitals.

Say we have a coefficient matrix $\boldsymbol{U}$ giving the expansion coefficients for our rotated orbitals in terms of the starting atomic orbitals. Then, the transformed integrals v_{pqrs} are obtained from the atomic orbital integrals $v^{AO}_{\mu\nu\kappa\lambda}$ through (assuming real coefficients, for simplicity)

$$v_{pqrs} = \sum_{\mu\nu\kappa\lambda} U_{p\mu} U_{q\nu} U_{r\kappa} U_{s\lambda} v^{\mathrm{AO}}_{\mu\nu\kappa\lambda} \tag{3.25}$$

As is well known, the four-index transformation should be carried out in four quarter-transformation steps corresponding to the four contractions with the coefficient matrices above. In our parallel transformation scheme, we consider the four steps in two stages; in the first stage we perform two quarter-transformations to construct half-transformed Coulomb and exchange intermediates J, K

$$J_{ab}(\nu, \kappa) = \sum_{\mu\lambda} U_{a\mu} U_{b\lambda} v^{\mathrm{AO}}_{\mu\nu\kappa\lambda} \tag{3.26}$$

$$K_{ab}(\nu, \kappa) = \sum_{\mu\lambda} U_{a\mu} U_{b\lambda} v^{\mathrm{AO}}_{\mu\nu\lambda\kappa} \tag{3.27}$$

while in the second stage, we perform the remaining quarter transformations on the J, K intermediates to obtain the final integrals

$$[J_{ab}]_{pq} = v_{apqb} = \sum_{\nu\kappa} J_{ab}(\nu, \kappa) U_{p\nu} U_{q\kappa} \tag{3.28}$$

$$[K_{ab}]_{pq} = v_{apbq} = \sum_{\nu\kappa} K_{ab}(\nu,\kappa) U_{p\nu} U_{q\kappa} \tag{3.29}$$

Note that for the purposes of optimizing the active orbitals, we only need the integrals that appear in the augmented Hessian. Thus, the ab indices in Eqns. 3.26 and 3.27 only need to run over the active orbitals while the pq indices need to run over all the closed-shell, active, and external orbitals.

In the first stage, we parallelize the construction of the J, K intermediates by dividing up the intermediates according to their untransformed AO indices. For example, the construction of $J_{ab}(\nu,\kappa)$ is divided amongst the processors according to the pair of indices (ν,κ); each processor is then responsible for constructing the J intermediates for all $(\bar{\nu},\bar{\kappa}) \in proc$. This allows us to also partition the AO integrals amongst the processors according to the same divided pair of indices $(\bar{\nu},\bar{\kappa})$; e.g. to construct $J_{ab}(\bar{\nu},\bar{\kappa})$ for $(\bar{\nu},\bar{\kappa}) \in proc$ we only need AO integrals such as $v^{AO}_{\mu\bar{\nu}\bar{\kappa}\lambda}$ for $(\bar{\nu},\bar{\kappa}) \in proc$ to be stored on that processor.

Once all J and K intermediates are constructed, we parallelize the second stage with respect to the transformed ab indices of the J, K intermediates. Thus ab is divided amongst the processors, and each processor constructs the final integrals $v_{\bar{a}pq\bar{b}}, v_{\bar{a}p\bar{b}q}$ for all $\{\bar{a}\bar{b}\} \in proc$. Since the first stage is parallelized over a pair of AO indices (ν,κ) (and the J and K intermediates are divided across the processors accordingly) while the second stage is parallelized over the two transformed indices (ab), we need to redistribute the intermediates J and K amongst the processors between the first and second stages. This is the main communication step.

In addition to above parallelization, further efficiencies can be gained by using the permutational and spatial symmetries of the integrals. Our complete parallelized algorithm, which uses these symmetries, is presented in pseudo-code

in Alg. (6) in the Appendix (A.2). The cost of the four-index integral transformation as implemented is $O((K^4k + K^3k^2)/n_p)$ for CPU, $O((K^4 + K^2k^2)/n_p)$ for disk space, $O(K^2k^2/n_p)$ for memory, and $O(K^2k^2)$ for overall communication, where K is the total number of orbitals, k is the number of active orbitals, and n_p is the number of processors.

To complete our efficient implementation of orbital optimization, we have also parallelized the remaining steps in the Augmented Hessian Newton-Raphson solver. These additional steps take up only a small part of the computational time and have an overall cost $O(K^2k^3/n_p)$ for CPU time, $O(K^2k^2/n_p)$ for memory, $O(Kk)$ for communication.

3.7 Localization schemes

Since DMRG is an inherently localized wavefunction ansatz, we need to localize the orbitals before we perform a DMRG calculation with them. The need for locality in DMRG algorithm becomes very apparent when we look at the connectivity in DMRG and the way we break up the system into site functions. For an efficient DMRG algorithm only nearby sites or orbitals should be strongly correlated and the correlations should decrease rapidly with distance. Thus, the active space orbitals have to be localized after each orbital rotation and the active integrals transformed into this local basis before each DMRG calculation.

There are a few commonly used localization methods :

- Boys localization.
- Edmiston-Ruedenberg localization.

- Pipek-Mezey localization.

Table 3.2: Characteristics and computational cost of various localization schemes.

Localization scheme	Criterion for localization	Cost scaling
Boys	Maximize distance between orbital centroids	$O(N^3)$
Edmiston-Ruedenberg	Maximize orbital self-repulsion energy	$O(N^5)$
Pipek-Mezey	Minimize number of atoms that an orbital spans	$O(N^3)$

Boys localization and Edmiston-Ruedenberg (ER) localization [14] have been the most widely used of the localization schemes and they have a very physically rational way of localizing the orbitals. In the Boys method the localization is done by maximizing the separation of orbital centroid and the ER method consists of maximizing the orbital self-repulsion energy.

Pipek-Mezey localization (PML) [15] scheme follows a different approach. It does not use the Coulombic properties as the criterion for localization. However using this technique the localized orbitals look quite similar to the ER localized orbitals.

In the PML scheme, a measure of delocalization is defined as

$$d_i = [\sum_{A=1}^{n} (Q_A^i)^2]^{-1} = [\sum_{A=1}^{n} (\sum_{\mu \in A} \sum_{\nu=1}^{m} C_{i\mu} C_{i\nu} S_{\mu\nu})^2]^{-1}, \tag{3.30}$$

where Q_A^i is the gross Mulliken population of orbital i on atom A, $C_{i\mu}$ is the MO coefficient of μth AO on ith MO and the $S_{\mu\nu}$ is the AO overlap between the

AOs μ and ν. The number of atoms is given as n. Thus, the d_i is the average number of atoms that the orbitals are delocalized on and serves as an efficient measure of delocalization. The PML scheme is to minimize this d_i. It shows better convergence than the Boys localization and has a cheaper computational cost scaling than ER localization.

In our work, we have used the Pipek-Mezey localization scheme.

3.8 Natural orbitals

In order to facilitate fast convergence in the orbital rotation step, natural orbitals are used. *Natural orbitals* are defined as the orbital basis obtained from the eigenfunctions of the spinless one particle reduced density matrix.

3.9 Complete orbital optimized DMRG-CASSCF algorithm

With the description of the density matrix evaluation in Sec. 3.5 and the orbital optimization and integral transformation in Sec. 3.6, we now have the basic ingredients to perform the DMRG-CASSCF algorithm, according to the general outline in Sec. 3.2.

The complete DMRG-CASSCF algorithm is as follows:

1. Localize the active space orbitals.
2. Transform the AO integrals to the active space basis and build the active space Hamiltonian.

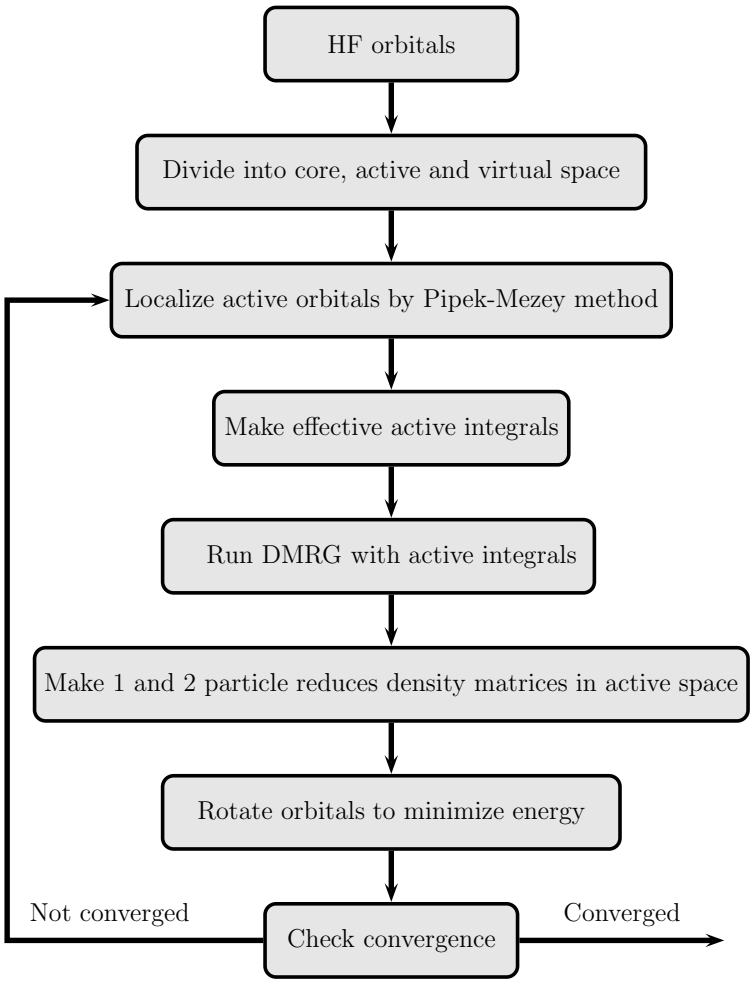


Figure 3.4: Flowchart of DMRG-CASSCF method. Starting with the Hartree-Fock orbitals (localized in the active space), the evaluation of DMRG wavefunction, reduced density matrix and orbital rotation are done self-consistently until the energy converges.

3. Perform the DMRG calculation using the active space Hamiltonian.

4. From the converged DMRG wavefunctions at each block configuration, assemble the one- and two-particle density matrices.

5. Using the density matrices, obtain the orbital gradient and orbital step from the Augmented Hessian Newton-Raphson solver.
6. From the orbital step, determine the new active space orbitals.
7. Goto 1. until convergence in the energy.

Steps 1.-6. constitute a single DMRG-CASSCF *macro-iteration*.

3.10 Conclusion

We have described how to efficiently implement orbital optimization using density matrix renormalization group wavefunction. By virtue of the compact nature of the DMRG wavefunction, this now enables us to handle much larger active spaces than are possible with traditional CASSCF algorithm.

BIBLIOGRAPHY

[1] D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, J. Chem. Phys. **128**, 144117 (2008).

[2] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **118**, 8551 (2003).

[3] G. K.-L. Chan, M. Kállay, and J. Gauss, J. Chem. Phys. **121**, 6110 (2004).

[4] J. Hachmann, W. Cardoen, and G. K.-L. Chan, J. Chem. Phys. **125**, 144101 (2006).

[5] J. Hachmann, J. J. Dorando, M. Avilés, and G. K.-L. Chan, J. Chem. Phys. **127**, 134309 (2007).

[6] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, Chem. Phys. **48**, 157 (1980).

[7] B. O. Roos, The complete active space self-consistent field method and its applications in electronic structure calculation, in *Advances in Chemical Physics; Ab Initio Methods in Quantum Chemistry II*, edited by Wiley, page 399, Chichester, 1987.

[8] P. J. Knowles and H.-J. Werner, J. Chem. Phys. **82**, 5053 (1985).

[9] D. L. Yeager, D. Lynch, J. Nichols, P. Jørgensen, and J. Olsen, J. Phys. Chem. **86**, 2140 (1982).

[10] B. H. Lengsfield III and B. Liu, J. Chem. Phys. **75**, 478 (1981).

[11] S. R. White and R. L. Martin, J. Chem. Phys. **110**, 4127 (1999).

[12] G. K.-L. Chan et al., Progress in Theoretical Chemistry and Physics **18**, 49 (2008).

[13] G. K.-L. Chan, J. Chem. Phys. **120**, 3172 (2004).

[14] J. W. Boughton and P. Pulay, J. Comput. Chem. **14**, 736 (1993).

[15] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

CHAPTER 4

APPLICATION OF DMRG-CASSCF TO POLYENES AND β-CAROTENE

(Some parts of this chapter are taken from Ref. [1].)

4.1 Introduction

Polyenes are the simplest conjugated systems, consisting of alternating singly and doubly bonded carbons arranged in a chain. They are valuable models not only to understand conjugated polymers of materials interest (e.g. polyacetylene is simply an infinite polyene) but also biological molecules such as the carotenoid and retinal families of pigments involved in photosynthesis and vision. These class of compounds are also important as they played a major role in the development of molecular orbital theory and are thus, important to electronic structure theory at its birth.

In these systems, the functionality of the molecules relies on the low-lying π-π^* excited states of the conjugated backbone, which serve as the conduits for energy transfer. The excited states are labeled by their symmetry under the C_{2h} point group, giving rise to A_g, B_g, A_u, B_u symmetry labels. Furthermore, they are usually given an additional $+/-$ label to indicate their approximate particle-hole symmetry. In Hamiltonians (such as the Hückel Hamiltonian) which support symmetric sets of energy states around the Fermi level, there is an additional symmetry associated with rotating the molecular orbital diagram so that the bonding and anti-bonding levels swap places [2]. Although particle-hole symmetry is not a true symmetry of the *ab-initio* electronic Hamiltonian, it is still customary to use such labels for the polyenes, in particular, because the

$+/-$ states have very different qualitative electronic structure; valence bond studies of the Hubbard model [3] show that the $+$ states consist mainly of ionic valence bond structures, while the $-$ states consist mainly of covalent valence bond structures [4–6].

4.2 Motivation

In our study of conjugated polymers using DMRG-CASSCF, we have looked at singlet states and henceforth we shall be considering singlet states only. The ground state of the polyenes is known to always be of A_g^- symmetry. The lowest dipole-allowed singlet transition, which has a predominantly HOMO→LUMO excitation character, has B_u^+ symmetry. However, contrary to what one might expect, this $1A_g^- \rightarrow 1B_u^+$ transition is not the lowest singlet transition [7, 8].

Rather, as shown by Aoyagi *et al*. in octatetraene [8], there is a *lower* dipole forbidden excitation, later identified as the $2A_g^-$ state, which can be rationalized in valence bond language as arising from a pair of singlet-triplet excitations in the two separate double bonds that recouple to form a singlet state [9–17]. Following the observation of the $2A_g^-$ state in octatetraene, there has been much debate over the correct ordering of the $2A_g^-$ and $1B_u^+$ excited states in the shorter polyenes, compounded both by experimental difficulties in observing the dipole-forbidden $2A_g^-$ state as well as theoretical challenges in achieving a balanced description of the two states, which are dominated by very different kinds of correlation, namely static correlation in the $2A_g^-$ state and dynamic correlation in the $1B_u^+$ state as also the correct degree of valence-Rydberg mixing that occurs in the very small polyenes, $\mathrm{C_4H_6}$ and $\mathrm{C_6H_8}$. Thus, despite repeated

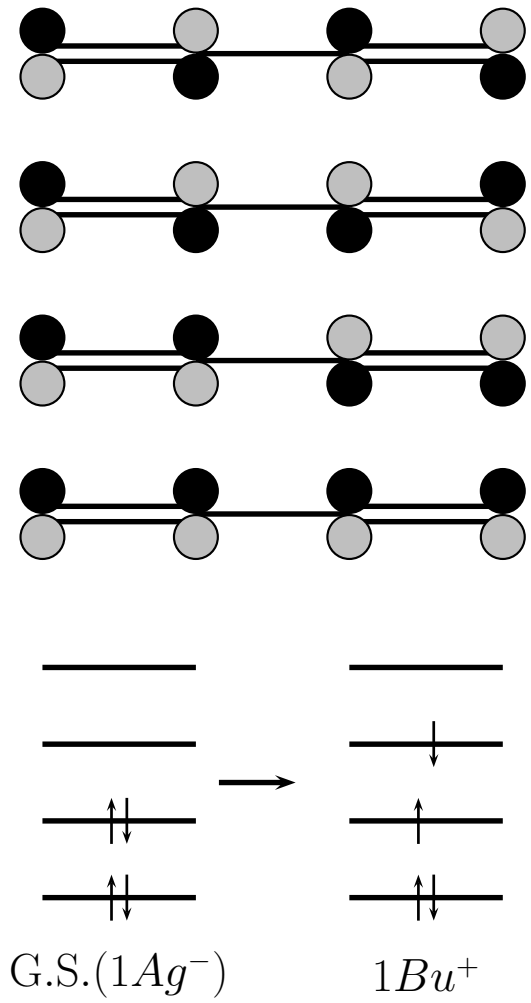


Figure 4.1: Ground state $1A_g^-$ and first optically active $1B_u^+$ excited state of C_4H_6.

reinvestigations since they were first observed 70 years ago [18, 19], the spectra of the polyenes have remained a mystery [20–24].

In longer polyenes and the biologically active carotenoid and retinal pigments, questions about the low-lying spectrum are not restricted simply to the $2A_g^-$ and $1B_u^+$ state ordering. Recent studies using resonance Raman excitation profiles (RREP) and electronic absorption spectroscopy on substituted polyenes in the

carotenoid family, have indicated the presence of additional dark states below the $1B_u^+$ state [25–29]. In particular, for the all-*trans*-carotenoids with (the number of double bonds) $n = 9 - 11$, Sashima *et al*. observed a $1B_u^-$ state between the $2A_g^-$ and $1B_u^+$ [26, 30]. More recently, Furuichi *et al*. observed a $3A_g^-$ level between the $1B_u^-$ and $1B_u^+$ states in carotenoids with $n = 11 - 13$, and assigned the tentative state ordering of $1A_g^- < 2A_g^-(S_1) < 1B_u^-(S_2) < 3A_g^-(S^*) < 1B_u^+(S_3)$ [29]. The assignment was made by extrapolating from the earlier PPP-MRDCI calculations by Tavan and Schulten on short polyenes ($n = 2 - 8$), which had predicted the existence of these additional states [6, 31].

To understand the long polyenes and carotenoids, there have been a number of density functional studies on them, but these calculations generally find that the low-lying state ordering in longer polyenes is dependent on the exchange-correlation functional [32–34].

To better understand the electronic structure of these low-lying states, the starting point would be to carry out an *ab-initio* multireference calculation, using the complete π-valence space. However, the large number of active π orbitals in the longer polyenes means that it is not possible to perform such calculations with traditional CAS algorithms for these systems. Hirao and coworkers [4, 35] carried out *incomplete valence* CASSCF and CASCI-MRMP using a (10,10) active space on the polyene series up to $C_{28}H_{30}$ and observed reasonable agreement with experiment. With our new orbital optimized DMRG-CASSCF procedure (described in Chapter III), we can now re-examine the low-lying excitations in these systems correlating the *complete* π- valence space even for the longer polyenes and carotenoids. By using a complete valence space as active orbitals, we can get a balanced and physically motivated active space.

4.3 Computational details

The polyene molecular geometries for $\mathrm{C_8H_{10}}, \mathrm{C_{12}H_{14}}, \mathrm{C_{16}H_{18}}, \mathrm{C_{20}H_{22}}, \mathrm{C_{24}H_{26}}$ were optimized at the density functional level using the B3LYP functional [36, 37] as implemented in GAUSSIAN03 [38]. The polyene molecules were constrained to have C_{2h} symmetry, with the C_2 axis as the z-axis. The cc-pVDZ basis [39] was used for all calculations.

In our DMRG-CASSCF calculations we used a complete π-valence space i.e. in $\mathrm{C_{24}H_{26}}$, this was a (24o, 24e) active space. To generate this active space, we first performed a restricted Hartree-Fock calculation in PSI3 [40, 41] to obtain canonical Hartree-Fock molecular orbitals.

From the RHF orbitals, it is not trivial to identify the appropriate π anti-bonding active orbitals because of significant $2p$-$3p$ mixing. Thus, we start by choosing the π bonding space which is quite straightforward as we can trivially choose $n/2$ orbitals from $n/2$ orbitals of A_u and B_g symmetries. Then, we constructed the anti-bonding component of the active space as a set of projected atomic orbitals (PAO), by first projecting out the π bonding space from a set of $2p_z$ atomic orbitals. These projected atomic orbitals were then symmetrically orthogonalized, then relocalized together with the bonding molecular orbitals, using the Pipek-Mezey procedure [42](details given in the Appendix A.1), to yield the complete active space in our calculations. The final set of active orbitals generated in this way resemble an orthogonal set of $2p_z$ orbitals.

Note that our initial active space does not correspond precisely to an active space obtained by selecting Hartree-Fock canonical orbitals. Thus DMRG energies obtained *before* orbital optimization do not correspond to typical CASCI

energies, but instead to CASCI energies obtained in our projected-atomic orbital (PAO) virtual space. This distinction is noted in our tables with the abbreviation DMRG-PAO-CASCI. After orbital optimization, however, our DMRG-CASSCF energies do correspond to true CASSCF energies, up to the accuracy of the DMRG calculation.

We carried out state-averaged DMRG-CASSCF calculations in the above active space with the one-site DMRG algorithm as described in Chapter II and III with a value of $M = 250$ and averaging over the 4 lowest eigenstates. The DMRG sweeps were converged to $10^{-10} E_h$ in the DMRG energy, which took roughly 30 DMRG sweeps. The number of renormalized states was increased smoothly from a starting value of $M = 50$ to the final value of $M = 250$. To aid the convergence of the DMRG sweeps in the one-site algorithm, we applied a system-environment perturbation as described in Ref. [43], with a starting magnitude of 10^{-3} that smoothly decreased to 0 after 20 sweeps. We estimate the remaining error in the DMRG energies at the $M = 250$ level from the exact FCI energies in the same active space to be less than $0.1\text{m}E_h$. Our DMRG calculations were combined with orbital rotation in a macro-iteration consisting of a converged DMRG calculation, an Augmented-Hessian step based orbital rotation, integral transformation, and orbital localization, as described in Sec. 3.9. Typically 10-15 macro-iterations of the complete DMRG/orbital optimization cycle were necessary to converge the energies to a tolerance of better than $10^{-6} E_h$. The convergence of the state energies with the number of macro-iterations is shown in Fig. 4.3.

The spatial and spin symmetries of excited states were assigned as follows. Firstly, all excited states were restricted to be of singlet spin symmetry through

the application of a shift $\lambda(\hat{S}^2 - \langle S\rangle(\langle S\rangle + 1))$ with $\lambda = 0.5$ [44]. To obtain the spatial symmetry, the ground state was assumed to be $1A_g^-$ as established by prior experimental and theoretical work. To determine whether the excited states were of A_g or B_u symmetry the transition dipole matrices were calculated between the states. Additionally, to determine the approximate particle-hole $+$ or $-$ symmetry we examined the magnitude of the transition dipoles; large transition dipoles for an allowed transition indicated that the transition involved a change of particle-hole symmetry between the states.

4.4 Results

4.4.1 Polyenes

Excitation energies

In Table 4.1 we present the energies, symmetries, and oscillator strengths for the ground state and first three excitations in the polyenes from $\mathrm{C_8H_{10}}$ to $\mathrm{C_{24}H_{26}}$. The DMRG-PAO-CASCI and DMRG-CASSCF entries for the $1A_g^-$ ground-states give the total energy in E_h; the other entries give the excitation energies from the ground state in eV. The estimated error of the DMRG-CASSCF energies from the exact CASSCF energies in the same active space is less than 0.1mE_h. The notation (n, m) denotes the active space used in the DMRG-PAO-CASCI and DMRG-CASSCF calculations. Oscillator strengths are in $a.u.$ for the ground-state, excited state transition. The CASCI-MRMP excitation energies are from Kurashige *et al.* [4]; note that these used at most a (10,10) active space. The

Table 4.1: Energies (in E_h for g.s. and eV for e.s.), symmetries, and oscillator strengths (in $a.u.$) for the lowest lying singlet excited states in conjugated polyenes.

Polyenes	Symmetry	DMRG PAO-CASCI	DMRG CASSCF	Oscillator Strength	CASCI-MRMP	Expt
C_8H_{10}	$1A_g^-$	-308.823021	-308.825879			
$(8,8)$	$2A_g^-$	6.33	4.69	Forbidden	4.26	3.54 [1]
	$1B_u^-$	7.49	5.88	0.0565	5.30	
	$3A_g^-$	7.95	6.60	Forbidden	7.20	
$C_{12}H_{14}$	$1A_g^-$	-462.661260	-462.670591			
$(12,12)$	$2A_g^-$	5.40	3.76	Forbidden	3.19	
	$1B_u^-$	6.30	4.74	0.0620	3.98	
	$3A_g^-$	7.01	5.59	Forbidden	5.12	
$C_{16}H_{18}$	$1A_g^-$	-616.499262	-616.514639			
$(16,16)$	$2A_g^-$	4.90	3.25	Forbidden	2.50	2.21 [2]
	$1B_u^-$	5.60	4.03	0.0502	3.10	
	$3A_g^-$	6.28	4.78	Forbidden	3.99	
$C_{20}H_{22}$	$1A_g^-$	-770.337112	-770.358327			
$(20,20)$	$2A_g^-$	4.60	2.93	Forbidden	2.04	(1.76) [3]
	$1B_u^-$	5.15	3.57	0.0427	2.51	(2.18) [3]
	$3A_g^-$	5.71	4.20	Forbidden	3.11	(2.47)[3]
$C_{24}H_{26}$	$1A_g^-$	-924.174795	-924.201821			
$(24,24)$	$2A_g^-$	4.42	2.73	Forbidden	1.70	(1.53)[3]
	$1B_u^-$	4.85	3.25	0.0384	2.05	(1.80)[3]
	$3A_g^-$	5.31	3.78	Forbidden	2.45	(2.02)[3]

[1] [45].
[2] [46].
[3] [29].

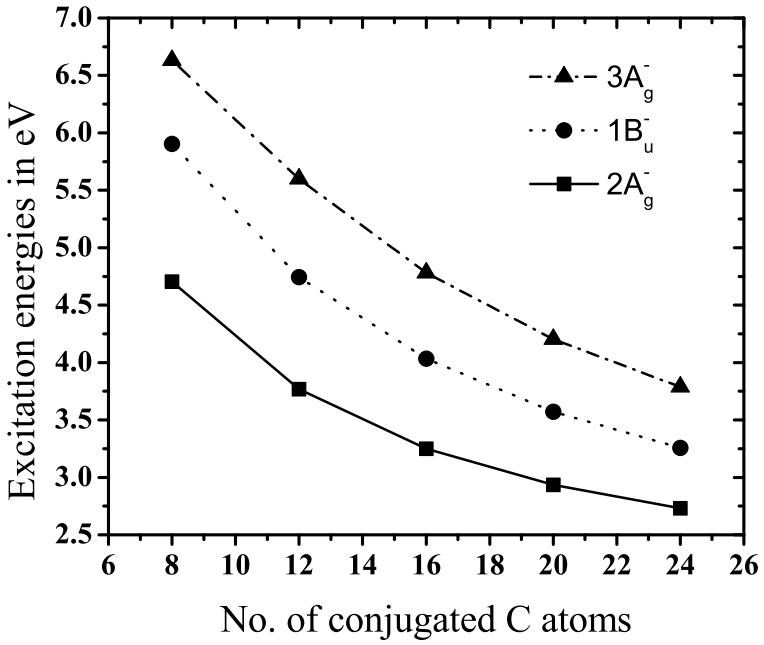


Figure 4.2: DMRG-CASSCF excitation energies in eV for the $2A_g^-$, $1B_u^-$ and $3A_g^-$ states in the conjugated polyenes $\mathrm{C_8H_{10}}$ to $\mathrm{C_{24}H_{26}}$.

experimental numbers in brackets are from measurements on the substituted polyene, spheroidene [29].

We have also plotted the DMRG-CASSCF excitation energies of the polyenes in Fig. 4.2.

We see that while our complete π-valence active space DMRG-CASSCF calculations generally over-estimate the excitation energies, they reproduce the correct experimental ordering of the lowest excited states with the exception of the missing $1B_u^+$ state (the HOMO-LUMO excitation), which should lie *below* the $3A_g^-$ in the shorter polyenes such as $\mathrm{C_8H_{10}}$. If we perform a state-averaged DMRG-CASSCF with 5 states in $\mathrm{C_8H_{10}}$, we find that the $1B_u^+$ state lies imme-

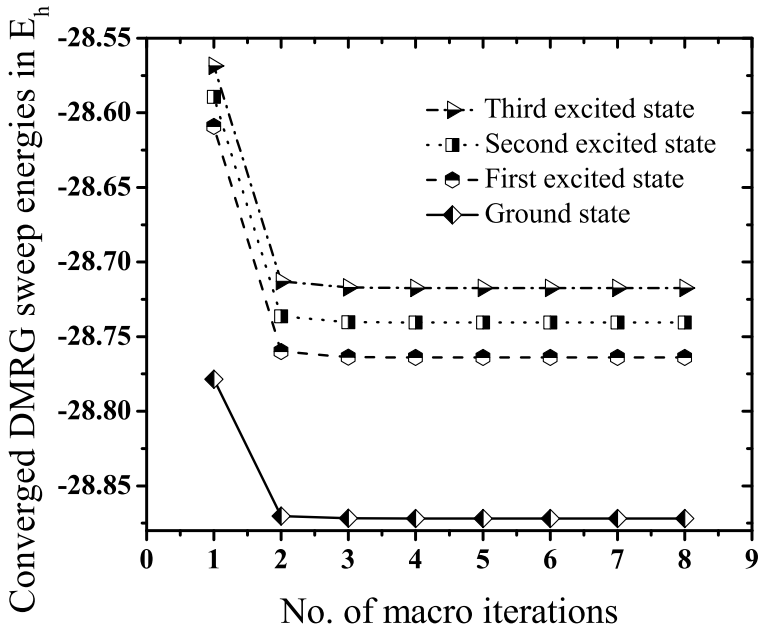


Figure 4.3: Converged DMRG sweep energies in Hartrees vs number of orbital optimization macro iterations in $C_{20}H_{22}$.

diately *above* the $3A_g^-$. This may seem strange given that CASSCF is generally believed to yield qualitatively correct electronic structure, but it reflects the wisdom from earlier studies on butadiene that σ-π correlation is very strong in the $1B_u^+$ state and must be included to obtain the correct balance between Rydberg and valence characters [10, 12, 47, 48]. Comparing with the calculations of Kurashige *et al.* [4], which despite having an incomplete valence active space include dynamic σ-π correlation through MRMP perturbation theory [49], further indicates that σ-π correlation would also lower the excitation energies of our other excited states.

From the variation of the CASSCF and the CASCI excitation energies we can see the effect of the relaxation energy for each excitation (see Fig. 4.3 and Table

4.1). The effect is least predominant in the ground state and most in the first excited $2^1A_g^-$ state. Thus the orbital rotation greatly improves the first excitation energy $1A_g^- \rightarrow 2A_g^-$. However, the $1A_g^- \rightarrow 1B_u^-$ is still greatly over-estimated, thus, indicating the need for inclusion of dynamic correlation.

Effect of active space

To better understand the effect of using a complete π valence space on the excitation energies, we have performed some small benchmark CASSCF calculations on $C_{12}H_{14}$ with $4-12$ active orbitals. These results are presented in Fig. 4.4.

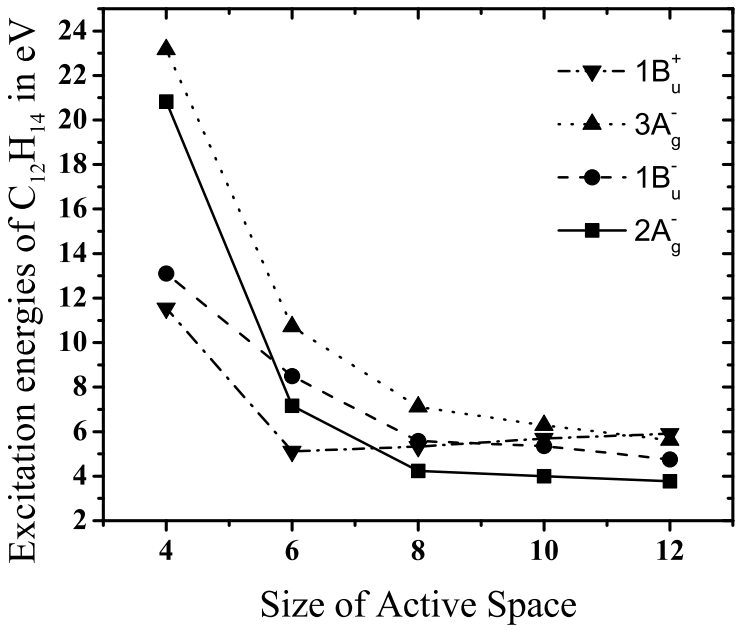


Figure 4.4: Change in CASSCF energies of the low-lying states of $C_{12}H_{14}$ as a function of increasing the active space from (4,4) to (12,12) (i.e. complete valence active space).

As can be seen, there is a very strong dependence of the excitation energies on

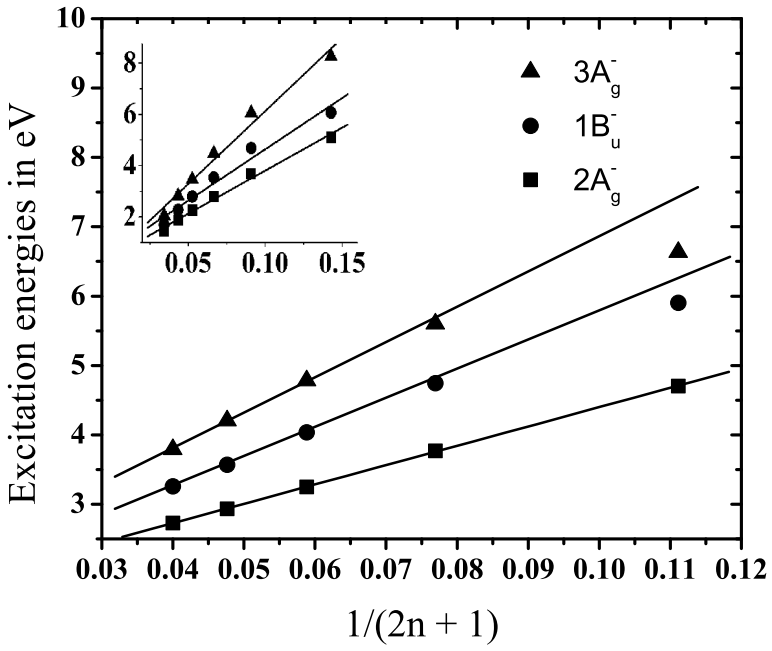


Figure 4.5: DMRG-CASSCF excitation energies for the low-lying singlet excited states of polyenes ranging from $C_{12}H_{14}$ to $C_{24}H_{26}$. The excitation energies are plotted against $1/(2n+1)$ where n is the number of double bonds. The ratio of the slopes for the different states are found to be 2:3.0:3.8 as compared to 2:3.1:3.8 experimentally. Inset: The same plot for the CASCI-MRMP energies from Kurashige *et al*. (Ref. [4]). As can be seen, these show a different and less linear dependence on $1/(2n+1)$.

the size of the active space, and even the order of the excitations changes. Thus, while an incomplete valence active space can yield an excited state ordering in better agreement with experiment, one is tempted to argue that it does not do so for the right reason.

Linear $1/(2n+1)$ fit

In Fig. 4.5, we plot our DMRG-CASSCF excitation energies as a function of the inverse chain length of the polyenes. Also shown (as an inset) is the same plot for the excitation energies obtained by Kurashige *et al*. [4]. It is easy to show that in a finite Hückel model with n sites, the excitation energies have a $\sin(k\pi/2(2n+1))$ chain length dependence, where k is a quasi-momentum number that labels the excitation. For long chains, this implies an asymptotic linear dependence on the inverse chain length $1/(2n+1)$. Tavan and Schulten conjectured that this asymptotic behavior held also in *interacting* systems, and presented evidence from MRD-CI calculations on short-chain Hubbard (n up to 7) and Pariser-Parr-Pople models (n up to 8) to support the conjecture [31]. The experimental resonance Raman excitation profiles from Sashima *et al*. [25] and Furuichi *et al*. [29] were also approximately fitted to the same inverse chain length behavior, although only over a small range of $n = 9 - 13$. We see from our results that while the $2A_g^-$ and $1B_u^-$ excitation energies fit the asymptotic $1/(2n+1)$ behavior well, the $3A_g^-$ state shows curvature more indicative of the sinusoidal dependence expected when $k \sim 2n+1$. This is consistent with interpreting the $3A_g^-$ as an excitation labeled by a larger quasi-momentum than $2A_g^-$. Interestingly, the excitation energies of Kurashige *et al*. show quite different chain-length dependence, with all three states showing much stronger curvature when their excitation energies are plotted against $1/(2n+1)$ in Fig. 4.5 (inlay). Fitting our excitation energies for $C_{16}H_{20}$, $C_{20}H_{24}$, $C_{24}H_{26}$ ($n = 8-12$) to the asymptotic dependence $1/(2n+1)$, we obtain slopes of $27.67eV$, $41.34eV$, $52.63eV$ for the $2A_g^-$, $1B_u^-$, $3A_g^-$ excitations, in reasonable agreement with the experimental slopes of $31.39eV$, $49.07eV$ and $59.63eV$. The ratio of the slopes for the different states is found to be $2 : 3.0 : 3.8$ as compared to $2 : 3.1 : 3.8$

experimentally.

Nature of the excited states

Table 4.2: Single particle nature of the polyene excitations (in %). For a given excited state (e.g. $2A_g^-$), the excitation weight of the transition $i \rightarrow j$ is given by $[\langle 1A_g^- | a_i^\dagger a_j | 2A_g^- \rangle]^2$. The total excitation weight is the sum of weights for all transitions; 100% indicates that the given excited state corresponds entirely to single excitations from the ground state. The transition labels $n \rightarrow m'$ are interpreted as follows: 1, 2, 3 ... denote HOMO, HOMO-1, HOMO-2 ... natural orbitals, while $1'$, $2'$, $3'$ denote LUMO, LUMO+1,LUMO+2 natural orbitals. As the polyenes increase in length, the total weight of the single excitations in the low-lying states becomes very small, $< 16\%$.

State	Excitation	No. of conjugated double bonds				
	weight	4	6	8	10	12
$2A_g^-$	$2 \rightarrow 1'$	10.9	8.6	6.6	5.3	4.3
	$1 \rightarrow 2'$	6.7	5.9	4.8	4.0	3.3
	Total	20.0	18.0	15.4	13.5	12.1
$1B_u^-$	$3 \rightarrow 1'$	14.5	10.2	7.9	6.3	5.2
	$1 \rightarrow 3'$	7.0	5.6	4.6	3.9	3.3
	Total	25.3	21.8	18.6	16.3	14.7
$3A_g^-$	$4 \rightarrow 1'$	21.3	12.8	9.3	7.1	5.6
	$1 \rightarrow 4'$	8.2	6.0	4.7	3.8	3.1
	Total	32.9	25.0	20.9	18.0	15.9

From the one particle transition density matrices we can analyze the single-particle character of our excitations. Given the density matrix element $w_{ij} =$

$\langle g.s.|a_i^\dagger a_j|excited\rangle$ where i, j are natural orbitals in the ground state, we define the weight of the $i \rightarrow j$ excitation as w_{ij}^2. The total single excitation weight is then $\sum_{ij} w_{ij}^2$. In Table 4.2 we give the largest excitation weights and the total single excitation weights for the low-lying polyene excited states as a function of the number of conjugated bonds. We see the $2A_g^-$, $1B_u^-$ and $3A_g^-$ states are dominated by many-particle excitations from the ground state (i.e. they have small single-particle excitation weights) and indeed the single-particle character of the excitations decreases even more as the chain-length increases. Remarkably, in $C_{24}H_{26}$ only $< 16\%$ of the excitation character of these states can be considered to be of a single-particle nature! These results are consistent with the analysis by Wormer and Dreuw using coupled cluster and propagator techniques [50].

Natural Orbital basis

Natural orbitals are defined as the orbitals in the basis of the eigenvectors $|\phi_i\rangle$ of the one particle reduced density matrix, ρ

$$\rho = \sum_i \rho_{ii} |\phi_i\rangle\langle\psi_i|. \tag{4.1}$$

In this basis, since the density matrix is diagonal, there are minimal interaction between the orbitals and thus, the orbital rotation is fastest converging in this basis. The natural orbitals of C_8H_{10} after the DMRG-CASSCF is converged are shown as a representative example.

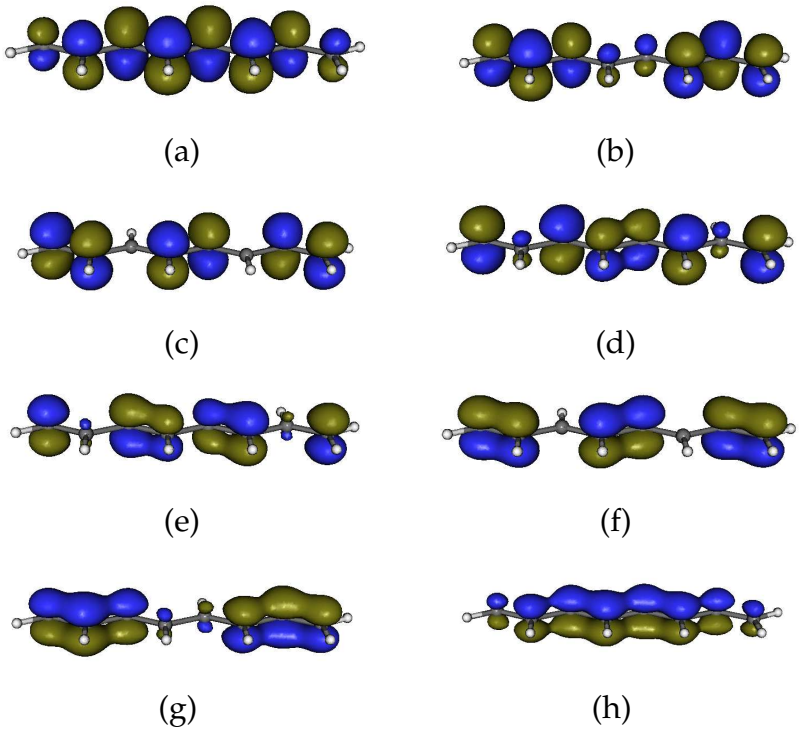


Figure 4.6: Natural Orbitals of C_8H_{10}.

4.4.2 Beta-Carotene

Carotenoids, the family of substituted polyenes, are the primary light harvesting pigments in the LH2 complex. Light harvesting proceeds by the absorption of visible light (in the green region) by the carotenoid molecules to form the singlet excited carotene (by the excitation of an electron from the ground state S_0 to optically active excited state S_2), followed by a fast internal conversion into an intermediate dark S_1 state from where the energy is transferred to the chlorophyll or bacteriochlorophyll molecule for photosynthesis. However, this rather simple three state model of light harvesting by carotenoid molecules was seriously questioned with the observation of various other excited states (S_x, S^* etc) [25–30] in between S_1 and S_2 states. Many essential questions remain

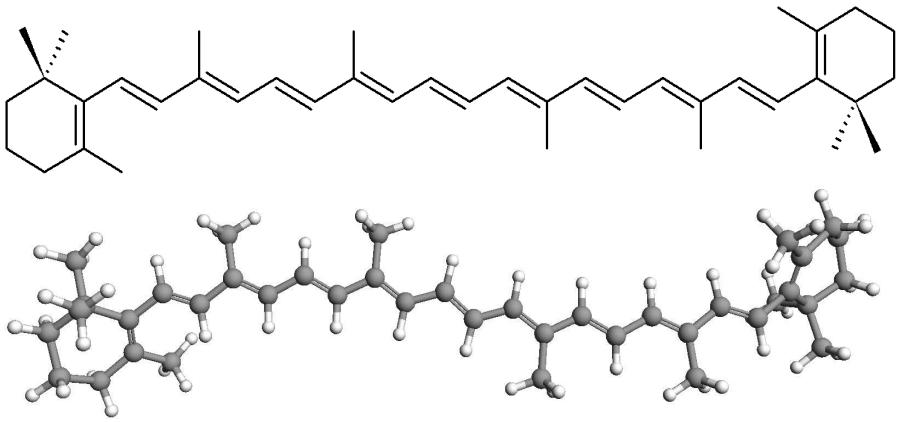


Figure 4.7: *s-cis* β-carotene.

unanswered as to the precise mechanism of this energy transfer [32–34, 50–53]. While the absorption of light places the carotenoid in the dipole allowed excited state, there can be a fast internal conversion to the aforementioned dark states of the polyene backbone, and thus multiple pathways for energy transfer to the bacteriochlorophyll.

We have chosen to study *s-cis* β-carotene (see Fig. 4.7) as a representative carotenoid. It is the dominant natural conformer although the *all-trans* form is also studied. Crystalline β-carotene has C_i symmetry with a conjugated backbone that lies almost entirely on the xy plane except for end groups which are twisted out of plane [54, 55]. (In the biological setting, carotenoid pigments usually adopt a twisted configuration in the conjugated backbone[56, 57]). There are 11 conjugated double bonds in the backbone. Our study employed the same calculation procedure as described in Sec. 4.3 with the exception that we used a 6-31G basis set in the DMRG-CASSCF calculation due to the large size of the molecule. State-averaged DMRG-CASSCF calculations were performed with 4 states and a (22,22) complete π-valence space, in the manner described in Sec. 4.3.

Table 4.3: DMRG-CASSCF energies, symmetries, and oscillator strengths for the lowest lying singlet excited states in β-carotene with the complete π-valence (22,22) active space. Total energies in E_h, excitation energies in eV, oscillator strengths in $a.u.$. The estimated error of the DMRG-CASSCF energies from the exact CASSCF energies in the same active space is less than 0.1mE_h. Oscillator strengths are for the ground-state, excited state transition.

Symmetry	DMRG-CASSCF total energy	Excitation energy	Oscillator Strength	Expt
$1A_g^-$	-1546.914545			
$2A_g^-$	-1546.804503	2.99	Forbidden	1.81 [1]
$1B_u^-$	-1546.781125	3.63	0.2025	2.05 [1]
$3A_g^-$	-1546.755822	4.31	Forbidden	(2.22) [2]

[1][26].
[2]Excitation measured for lycopene [29].

In Table 4.3 we present the energies, symmetries, and oscillator strengths for the ground state and first three excitations in β-carotene. We reproduce the state ordering $1A_g^- < 2A_g^- < 1B_u^- < 3A_g^-$ as assigned by Furuichi *et al*. [29] (note that the $1B_u^+$ which does not appear in our calculation indeed lies above the $3A_g^-$ state in this molecule). However, just as in the polyenes, the excitation energies from the DMRG-CASSCF procedure are generally over-estimated in comparison with experiment, most likely due to the lack of σ-π dynamic correlation.

A question that has received some attention in the literature is the effective conjugation length of carotenoids, since the presence of substituents and non-planar geometries are expected to modify this from the naive value deduced from the Lewis structure [58]. Formally, β-carotene has 11 double bonds in the

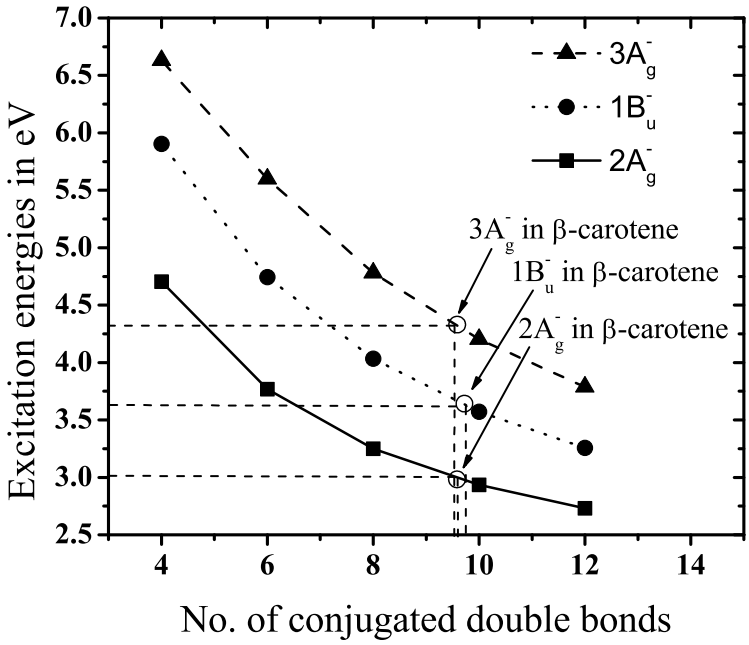
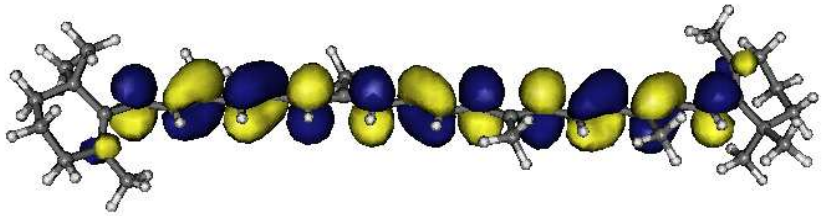
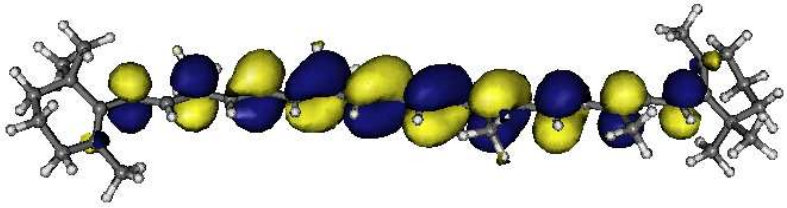


Figure 4.8: Polyene and carotene excitation energies vs. number of double bonds: the β-carotene excitation energies when fitted to the polyene excitation energies give an effective conjugation length of $9.5 - 9.7$.

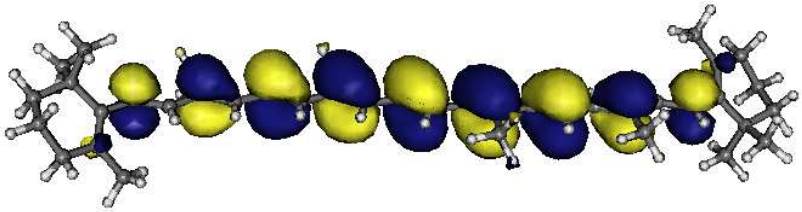
polyene backbone, but by comparing the excitation energies of the polyenes with our β-carotene excitation energies, we can estimate a reduced conjugation length of 9.5-9.7 bonds, which is very close to the experimental estimate of 9.7 of Onaka *et al*. [28]. This reduced conjugation length results from the twist in the carotene end-groups. In Fig. 4.9 we plot the DMRG-CASSCF natural orbitals corresponding to the HOMO, HOMO-1, LUMO, and LUMO+1. As can be seen, there is very little density in these orbitals on the carotene end-groups, and this is consistent with our reduced effective conjugation length.



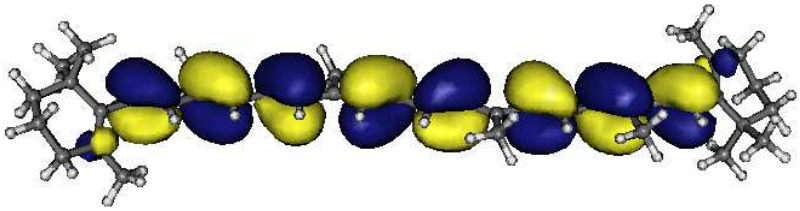
(a) Lowest Unoccupied Natural Orbital + 1



(b) Lowest Unoccupied Natural Orbital



(c) Highest Occupied Natural Orbital



(d) Highest Occupied Natural Orbital -1

Figure 4.9: Natural orbitals of β carotene corresponding to the HOMO-1 through LUMO+1 states. These orbitals participate in the lowest-lying singlet excitations in β carotene and contain little density on the nonplanar end groups.

4.5 Conclusion

In this chapter, we have described the application of DMRG-CASSCF to low-lying excitations of polyenes from $\mathrm{C_8H_{10}}$ to $\mathrm{C_{24}H_{26}}$ as well as light harvesting pigment β-carotene. Thus, we have used DMRG-CASSCF with up to (24,24) active space. The calculations have reproduced the state ordering of the dark states that have been recently observed by resonance Raman studies. However, since we have not considered the dynamic $\sigma - \pi$ correlation which plays a significant role in the optically active HOMO-LUMO $1B_u^+$ transition, the excitation energy for this state is overestimated.

With the application of DMRG-CASSCF to linear polyenes and carotenoids, we conclude the static correlation part of the thesis. To include the dynamic correlation, we will consider perturbative as well as canonical transformation methods in the next few chapters.

BIBLIOGRAPHY

[1] D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, J. Chem. Phys. **128**, 144117 (2008).

[2] R. Pariser, J. Chem. Phys. **24**, 250 (1956).

[3] J. Hubbard, Proc. R. Soc. London, Ser. A **276**, 238 (1963).

[4] Y. Kurashige, H. Nakano, Y.Nakao, and K. Hirao, Chem. Phys. Lett. **400**, 425 (2004).

[5] S. Ramasesha, S. K. Pati, H. R. Krishnamurthy, Z. Shuai, and J. Brédas, Phys. Rev. B **54**, 7598 (1996).

[6] P. Tavan and K. Schulten, J. Chem. Phys. **85**, 6602 (1986).

[7] B. S. Hudson and B. E. Kohler, J. Chem. Phys. **59**, 4984 (1973).

[8] M. Aoyagi, I. Ohmine, and B. E. Kohler, J. Chem. Phys. **94**, 3922 (1990).

[9] L. Serrano-Andrés, J. Sanchez-Mańn, and I. Nebot-Gil, J. Chem. Phys. **97**, 7499 (1992).

[10] R. P. Hosteny, T. H. Dunning Jr., R. R. Gilman, A. Pipano, and I. Shavitt, J. Chem. Phys. **62**, 4764 (1975).

[11] R. J. Cave, J. Chem. Phys. **92**, 2450 (1990).

[12] R. J. Cave and E. R. Davidson, J. Phys. Chem. **92**, 614 (1988).

[13] B. R. Brooks and H. F. Schaefer III, J. Chem. Phys. **68**, 4839 (1978).

[14] C. Petrongolo, R. J. Buenker, and S. D. Peyerimhoff, J. Chem. Phys. **76**, 3655 (1982).

[15] J. Lappe and R. J. Cave, J. Phys. Chem. A **104**, 2294 (2000).

[16] A. C. Lasaga, R. J. Aerni, and M. Karplus, J. Chem. Phys. **73**, 5230 (1980).

[17] V. Bachler and K. Schaffner, Chem. Eur. J. **6**, 959 (2000).

[18] J. Stark and P. Lipp, Physik. Chem. **86**, 36 (1913).

[19] W. C. Price and A. D. Walsh, Proc. Roy. Soc. London **174A**, 220 (1940).

[20] A. Sabljić and R. McDiarmid, J. Chem. Phys. **84**, 2062 (1986).

[21] R. McDiarmid and A.-H. Sheybani, J. Chem. Phys. **89**, 1255 (1988).

[22] D. G. Leopold, R. D. Pendley, J. L. Roebber, R. J. Hemley, and V. Vaida, J. Chem. Phys. **81**, 4218 (1984).

[23] J. P. Doering and R. McDiarmid, J. Chem. Phys. **73**, 3617 (1980).

[24] K. L. D'Amico, C. Manos, and R. L. Christensen, J. Am. Chem. Soc. **102**, 1777 (1980).

[25] T. Sashima, H. Nagae, M. Kuki, and Y. Koyama, Chem. Phys. Lett. **299**, 187 (1999).

[26] T. Sashima, Y. Koyama, T. Yamada, and H. Hashimoto, J. Phys. Chem. B **104**, 5011 (2000).

[27] R. Fujii et al., J. Phys. Chem. A **105**, 5348 (2001).

[28] K. Onaka et al., Chem. Phys. Lett. **315**, 75 (1999).

[29] K. Furuichi, T. Sashima, and Y. Koyama, Chem. Phys. Lett. **356**, 547 (2002).

[30] G. Cerullo et al., Science **298**, 2395 (2002).

[31] P. Tavan and K. Schulten, Phys. Rev. B **36**, 4337 (1987).

[32] P. J. Walla, P. A. Linden, C.-P. Hsu, G. D. Scholes, and G. R. Fleming, Proc. Nat. Acad. Sci. **97**, 10808 (2000).

[33] C.-P. Hsu, P. J. Walla, M. Head-Gordon, and G. R. Fleming, J. Phys. Chem. B **105**, 11016 (2001).

[34] C.-P. Hsu, S. Hirata, and M. Head-Gordon, J. Phys. Chem. A **105**, 451 (2001).

[35] K. Nakayama, H. Nakano, and K.Hirao, Int. J. Quantum Chem. **66**, 157 (1998).

[36] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[37] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[38] M. J. Frisch et al., GAUSSIAN 03, REVISION C.02, Gaussian, Inc., Wallingford CT, 2004, see http://www.gaussian.com/.

[39] T. H. Dunning Jr., J. Chem.Phys. **90**, 1007 (1989).

[40] T. D. Crawford et al., PSI 3.2 (2003), see www.psicode.org.

[41] T. D. Crawford et al., J. Comput. Chem. **28**, 1610 (2007).

[42] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

[43] S. R. White, Phys. Rev. B. **72**, 180403 (2005).

[44] G. Moritz, A. Wolf, and M. Reiher, J. Chem. Phys. **123**, 184105 (2005).

[45] M. F. Granville, G. R. Holtom, B. E. Kohler, R. L. Christensen, and K. D'Amico, J. Chem. Phys. **70**, 593 (1979).

[46] B. E. Kohler, C. Spangler, and C. Westerfield, J. Chem. Phys. **89**, 5422 (1988).

[47] L. Serrano-Andrés, M. Merchán, I. Nebot-Gil, R. Lindh, and B. O. Roos, J. Chem. Phys. **98**, 3151 (1993).

[48] R. Lindh and B. O. Roos, Int. J. Quantum Chem. **XXXV**, 813 (1989).

[49] K. Hirao, Chem. Phys. Lett. **190**, 374 (1992).

[50] A. Dreuw, G. Fleming, and M. Head-Gordon, Phys. Chem. Chem. Phys. **5**, 3247 (2003).

[51] V. Sundström, Progress in Quantum Electronics **24**, 187 (2000).

[52] A. Damjanovic, T. Ritz, and K. Schulten, Phys. Rev. E **59**, 3293 (1999).

[53] A. Dreuw, G. R. Fleming, and M. Head-Gordon, Phys. Chem. Chem. Phys. **5**, 3247 (2003).

[54] S. Schlücker, A. Szeghalmi, M. Schmitt, J. Popp, and W. Kiefer, Journal of Raman Spectroscopy **34**, 413 (2003).

[55] K. V. Berezin and V. V. Nechaev, Journal of Applied Spectroscopy **72**, 164 (2005).

[56] Y. Wang, L. Mao, and X. Hu, Biophysical Journal **86**, 3097 (2004).

[57] P. Qian et al., Photochemistry and Photobiology **74**, 444 (2001).

[58] H. A. Frank, Archives of Biochemistry and Biophysics **385**, 53 (2001).

CHAPTER 5

DYNAMIC CORRELATION - CUMULANT APPROXIMATED N-ELECTRON VALENCE PERTURBATION THEORY

(Some parts of this chapter are taken from Ref. [1].)

5.1 Dynamic correlation

Dynamic correlation, as defined in Chapter I, is the electronic correlation that arises due to the cusp like nature of the wavefunction at the limit of electron electron coalescence. Therefore, it is very slowly convergent with the number of basis functions. Dynamic correlation is generally a quantitative correction to the qualitatively correct static correlated wave-function, i.e. the effect of dynamic correlation to the chemically important correlations is relatively lower than the effect of static correlation. However, the inclusion of dynamic correlation is crucial to obtain physical observables with experimental accuracy.

The dynamic correlation methods that can be used for multireference problems can mainly be divided into three groups - perturbative methods, multireference configuration interaction methods and the multireference coupled cluster (MRCC) [2–4] methods.

5.2 Cumulant approximated n-electron valence perturbation theory

Among the dynamic correlation methods for multireference problem, the configurational interaction methods are the most accurate but very expensive computationally. The MRCC methods lack the simple structure and accuracy of the single reference coupled cluster methods.

Thus, considering that the dynamic correlation is generally a small quantitative correction (to the chemically relevant observables) and the need for a lower computationally scaling method, perturbation theory, even at the second-order level, can be viewed as a practical and easiest tool to handle this kind of correlation. Perturbation theory on top of an arbitrary set of configurations in the valence space is known as multireference perturbation theory. The most popular forms today are the (essentially equivalent) complete active space second order perturbation theory (CASPT2) [5] and multireference Møller Plesset perturbation theory (MRMP) [6, 7] and more recently, n-electron valence perturbation theory (NEVPT2) [8].

Despite the many predictive successes of multireference perturbation theory, such methods face a number of limitations. A severe one is the high computational cost as a function of the number of active orbitals, which prohibits their application to large molecules with many active orbitals. For example, current implementations of NEVPT2 are limited to 14 active orbitals, but much larger active spaces are accessible through methods such as the density matrix renormalization group, [9–16], restricted-active-space self-consistent field, [17], or generalized valence bond methods [18]. Improved algorithms (in particu-

lar, efficient choices of when to employ internal contraction [19]) ameliorate but do not remove the sharp increase in complexity of the multireference perturbation theory as the number of active orbitals and determinantal configurations increases.

In this chapter, we explore the possibility of constructing approximate forms of multireference perturbation theory which are not at the outset limited to small active spaces. We take as our starting point the n-electron valence perturbation theory in its strongly contracted form (SC-NEVPT2), which is one of the simplest (and yet remarkably successful) multireference perturbation theories available [20]. The basic idea of the work is quite simple. The complexity of multireference perturbation theories can be understood to arise from their dependence on contributions from high-order (three- and four-body) density matrices involving the active orbitals. We will explore the possibility of approximating these density matrices via the one- and two-body density matrices using cumulant type expansions, removing any dependence on any more complex quantities. Naturally, this will introduce a degree of error, and our aim is to establish how tolerable such an error really is. To be successful, the error introduced by such cumulant type approximations should not be larger than the intrinsic error of second-order multireference perturbation theory.

Cumulant approximations have been employed by many different workers in electronic structure [21–25]. In earlier work from our group, cumulant and operator decompositions have played a role in the formulation of the canonical transformation method, an exponential based description of dynamical correlation for multireference problems [26–28]. Unlike the stated objective of reduced density matrix methods, we are not trying here to use the cumulants as the pri-

mary variational descriptors of the electronic structure; instead, they are introduced as computational approximations to quantities that arise naturally in the multireference perturbation theory. Nonetheless, while our goals and presentation are different from some works in the reduced density matrix area, some differences can be viewed as simply matters of terminology and philosophy (e.g. is multireference perturbation theory to be described as a four-body density matrix energy functional - to be approximated as a two-body density matrix energy functional, or as an internally contracted wavefunction based method where cumulants are used to approximate certain intermediates) and so there are many natural connections.

5.3 Theory

5.3.1 Strongly contracted n-electron valence perturbation theory

We present a brief review of second order n-electron valence perturbation theory (NEVPT2). Our presentation closely follows that given in Ref. [8, 29].

NEVPT2 is a second-order Rayleigh-Schrödinger perturbation theory, which differs from other kinds of multireference perturbation theory such as CASPT2 in its choice of zeroth order Hamiltonian and the representation of the first order wavefunction. Consider the zeroth order wavefunction $\Psi^{[0]}$ (with zeroth order energy $E^{[0]}$) which is the solution of some two-electron Hamiltonian H_{act} in a CASCI problem (typically obtained through a CASSCF calculation). We first es-

tablish some notations. We shall refer to core orbitals (doubly occupied in all CASCI configurations), active orbitals (variable occupancy in all CASCI configurations) and virtual orbitals (unoccupied in all CASCI configurations) (see Chapter I). Core orbitals will be associated with labels $i, j, k \ldots$, active orbitals with labels $a, b, c \ldots$, and virtual orbitals with labels $r, s, t \ldots$. The zeroth order Hamiltonian in NEVPT2 is the Dyall Hamiltonian [30], which consists of Fock type operators in the core and virtual spaces and the CASCI Hamiltonian H_{act} in the active space, i.e.

$$H^{[0]} = f_{\text{core}} + H_{\text{act}} + f_{\text{ext}} \tag{5.1}$$

$$f_{\text{core}} = \sum_i \epsilon_i a_i^\dagger a_i + C$$

$$H_{\text{act}} = \sum_{ab \in \text{act}} h_{ab}^{\text{eff}} a_a^\dagger a_b + \sum_{abcd \in \text{act}} v_{abcd} a_a^\dagger a_b^\dagger a_c a_d$$

$$f_{\text{ext}} = \sum_r \epsilon_r a_r^\dagger a_r \tag{5.2}$$

$$H^{[0]} \Psi^{[0]} = E^{[0]} \Psi^{[0]} \tag{5.3}$$

(Note that the one-body part of H_{act} includes the Coulomb field from the core electrons, i.e. $h_{ab}^{\text{eff}} = h_{ab} + \sum_{i \in \text{core}} (ii|ab) - (ia|ib)$ and C is a constant $2 \sum_{i \in \text{core}} h_{ii} + \sum_{ij \in \text{core}} 2(ii|jj) - (ij|ij) - \sum_{i \in \text{core}} \epsilon_i$ chosen so that the expectation value of the Dyall Hamiltonian with $\Psi^{[0]}$ is the CASCI energy. The core and virtual orbitals are taken to be canonical orbitals of Fock operators defined using CASCI one-particle density matrices).

From $H^{[0]}$, we define a Rayleigh-Schrödinger series in terms of the perturbation $V = H - H^{[0]}$. It is seen that $E^{[1]}$ vanishes since $\langle \Psi^{[0]} | H | \Psi^{[0]} \rangle = \langle \Psi^{[0]} | H^{[0]} | \Psi^{[0]} \rangle$. The first non-vanishing contribution to the energy is thus at 2nd order, which requires the first-order correction to the wavefunction $\Psi^{[1]}$. In the so-called *strongly contracted* NEVPT2 (SC-NEVPT2) which is of interest here, $\Psi^{[1]}$ is ex-

panded in terms of a *highly restricted* set of "perturber" functions. These are classified into eight spaces, $(0), (+1), (-1), (+2), (-2), (+1)', (-1)', (0)'$ which differ by the pattern of excitations involving core and virtual orbitals, and the number of particles or holes introduced into the active space. (The number $(+1)$ for example, denotes 1 particle introduced into the active space. Note also that we are using round brackets to label these eight spaces, whereas square brackets are used to denote orders of perturbation theory).

The perturber functions in each of the eight spaces are contracted sets of determinants, where the contraction coefficients are defined from the perturbation V. We first divide V into eight components which connect the reference wavefunction and the eight different spaces

$$\begin{aligned} V = & \sum_{i<j,r<s} V^{(0)}_{ijrs} + \sum_{i<j,r} V^{(+1)}_{ijr} \\ & + \sum_{r<s,i} V^{(-1)}_{rsi} + \sum_{i<j} V^{(+2)}_{ij} \\ & + \sum_{r<s} V^{(-2)}_{rs} + \sum_{i} V^{(+1)'}_{i} \\ & + \sum_{r} V^{(-1)'}_{r} + \sum_{ir} V^{(0)'}_{ir}, \end{aligned} \tag{5.4}$$

where the eight component perturbations are defined through (using the com-

pact second quantized notation e.g. $a_{ij}^{rs} = a_r^\dagger a_s^\dagger a_i a_j$),

$$
\begin{aligned}
V_{ijrs}^{(0)} &= \langle rs||ij\rangle a_{ij}^{rs} \\
V_{ijr}^{(+1)} &= \sum_a \langle ra||ji\rangle a_{ij}^{ra} \\
V_{rsi}^{(-1)} &= \sum_a \langle rs||ia\rangle a_{ai}^{rs} \\
V_{ij}^{(+2)} &= \sum_{a<b} \langle ab||ji\rangle a_{ij}^{ab} \\
V_{rs}^{(-2)} &= \sum_{a<b} \langle rs||ba\rangle a_{ab}^{rs} \\
V_i^{(+1)'} &= \sum_{a<b,c} \langle ab||ic\rangle a_{ci}^{ab} + \sum_{aj} \langle aj||ij\rangle a_{ji}^{aj} + \sum_a \langle a|h|i\rangle a_i^a \\
V_r^{(-1)'} &= \sum_{a,b<c} \langle ra||cb\rangle a_{bc}^{ra} + \sum_{aj} \langle rj||aj\rangle a_{ja}^{rj} + \sum_a \langle r|h|a\rangle a_a^r \\
V_{ir}^{(0)'} &= \sum_{ab} \langle ra||ib\rangle a_{bi}^{ra} + \sum_j \langle rj||ij\rangle a_{ji}^{rj} + \langle r|h|i\rangle a_i^r.
\end{aligned} \tag{5.5}
$$

Note that orbital labels of the component perturbations refer to only inactive (core or virtual) orbitals; all active indices are summed over.

The perturber functions are then generated by applying each of the component perturbations to the zeroth order wavefunction and normalizing. Thus the eight classes of perturber functions are

$$
\begin{aligned}
\Phi_{ijrs}^{(0)} &= \frac{1}{\sqrt{N_{ijrs}^{(0)}}} V_{ijrs}^{(0)} \Psi^{[0]} \\
\Phi_{ijr}^{(+1)} &= \frac{1}{\sqrt{N_{ijr}^{(+1)}}} V_{ijr}^{(+1)} \Psi^{[0]} \\
&\dots
\end{aligned} \tag{5.6}
$$

where we have introduced the norm e.g. $N_{ijr}^{(+1)} = \langle \Psi^{[0]} | V_{ijr}^{(+1)\dagger} V_{ijr}^{(+1)} | \Psi^{[0]} \rangle$. Note that every perturber function is orthogonal to every other perturber function. Consequently, in contrast to internally contracted CASPT2 or partially con-

tracted NEVPT2, there is no need for diagonalization of the overlap matrix between perturber functions. This feature of the strongly contracted theory is advantageous if we want to perform calculations with a large active space (e.g. with DMRG or another method), where the overlap diagonalization may prove to be a bottleneck.

Now that we have defined the perturber functions, we can expand the first order wavefunction in terms of this restricted basis

$$\Psi^{[1]} = \sum_{i<j,r<s} c^{(0)}_{ijrs} \Phi^{(0)}_{ijrs} + \sum_{i<j,r} c_{ijr} \Phi^{(+1)}_{ijr} + \dots \tag{5.7}$$

This defines the *strongly-contracted* NEVPT2 approximation to the first order wavefunction. The coefficients and energy contribution can be obtained using the the standard Rayleigh-Schrödinger expressions, evaluated in sum-over-states form. Taking the $(+1)$ subspace as an example, (and dropping the $(+1)$ labels below for convenience), we have

$$\begin{aligned} c_{ijr} &= -\frac{\langle \Phi_{ijr} | V | \Psi \rangle}{E_{ijr} - E^{[0]}} \\ &= -\frac{\langle \Phi^{(+1)}_{ijr} | V^{(+1)}_{ijr} | \Psi \rangle}{E_{ijr} - E^{[0]}} \\ &= -\frac{\sqrt{N_{ijr}}}{E_{ijr} - E^{[0]}} \end{aligned} \tag{5.8}$$

$$E^{[2]} = -\sum_{i<j,r} \frac{N_{ijr}}{E_{ijr} - E^{[0]}} \tag{5.9}$$

where E_{ijr} is the zeroth order energy of the perturber function i.e. $\langle \Phi^{(+1)}_{ijr} | H^{[0]} | \Phi^{(+1)}_{ijr} \rangle$.

Table 5.1: Highest rank reduced density matrices appearing in the energy contributions for the eight subspaces of NEVPT2. The energy expression involves a numerator and denominator (see Eqn. 5.9); the highest rank reduced density matrices contributing to the numerator and denominator separately are shown in columns 3 and 4. Active space density matrices do not contribute to the energy of the (0) space.

Subspace	All	Num.	Denom.
(0)	n.a.	n.a.	n.a.
$(+1)$	2	1	2
(-1)	2	1	2
$(+2)$	3	2	3
(-2)	3	2	3
$(+1)'$	4	3	4
$(-1)'$	4	3	4
$(0)'$	3	2	3

5.3.2 Cumulant approximated strongly contracted NEVPT2

Evaluating matrix elements in NEVPT2 for the coefficients and for the energy contributions is not a simple matter computationally. This can be understood because matrix elements involving perturber functions that allow active orbital relaxation (i.e. semi-internal type excitations in $(+1)', (-1)'$ spaces, of the form $a^{rc}_{ab}|\Psi\rangle$) involve long strings of active orbital operators. We can examine the complexity of NEVPT2 by reducing all matrix element expressions to traces of reduced density matrices with appropriate integrals. Depending on the subspace in NEVPT2, different density matrices are involved (see Table 5.1) but in the worst case (for the $(-1)', (+1)'$ subspaces), just as in other multireference

perturbation theories such as CASPT2, four-particle reduced density matrices formally appear. This greatly increases the complexity of these multireference perturbation theories relative to the single-reference counterpart, Møller-Plesset theory. For example, the computational scaling of the NEVPT2 implementation described in Ref. [29] is $O(n_{act}^9)$ to construct the intermediate expressions in subspaces $(+1)', (-1)'$. This presents a fundamental limitation if we wish to use the NEVPT2 method in conjunction with a reference function obtained in a large active space. As noted in the early articles on NEVPT2 [8], the theory is general and can in principle be combined with non-CASCI references (such as a GVB reference, or a DMRG reference) which are not limited to the small active spaces of CASCI.

One way to reduce the complexity of NEVPT2 theory (and multireference perturbation theories in general) is to remove the explicit or implicit dependence of the energy expressions on high particle density matrices. (By implicit dependence, we refer to such algorithms, often used in CASPT2 implementations, where the higher particle density matrices are not explicitly constructed, and instead where their contributions are computed directly on the fly using the determinantal expansion of the CASCI wavefunction). Ideally we would hope that no quantities more complex than the one- and two-particle active space density matrices should appear. This is achieved in the following cumulant approximated strongly contracted NEVPT2 (cu-SC-NEVPT2) and cumulant with diagonals approximated strongly contracted NEVPT2 (cud-SC-NEVPT2) methods.

cu-SC-NEVPT2: In the cu-SC-NEVPT2 approximation, we use cumulant approximations of the 3- and 4-particle density matrices in terms of the 1- and

2-particle density matrices. For the density matrices, we use the notation $\gamma^{abc}_{def} = \langle a^{abc}_{def} \rangle$. These approximations are then given as

$$\begin{aligned}
\gamma^{abc}_{def} \Rightarrow & \sum (-1)^x \gamma^a_d \gamma^{bc}_{ef} (9 \text{ terms}) - 2 \sum (-1)^x \gamma^a_d \gamma^b_e \gamma^c_f (6 \text{ terms}) \\
&= \gamma^a_d \gamma^{bc}_{ef} - \gamma^a_e \gamma^{bc}_{df} + \gamma^a_f \gamma^{bc}_{de} \\
&- \gamma^b_d \gamma^{ac}_{ef} + \gamma^b_e \gamma^{ac}_{df} - \gamma^b_f \gamma^{ac}_{de} + \gamma^c_d \gamma^{ac}_{ef} - \gamma^c_e \gamma^{ac}_{df} + \gamma^c_f \gamma^{ac}_{de} \\
&- 2(\gamma^a_d \gamma^b_e \gamma^c_f - \gamma^a_d \gamma^c_e \gamma^b_f + \gamma^b_d \gamma^c_e \gamma^a_f - \gamma^b_d \gamma^a_e \gamma^c_f \\
&+ \gamma^c_d \gamma^a_e \gamma^b_f - \gamma^c_d \gamma^b_e \gamma^a_f) \qquad (5.10) \\
\gamma^{abcd}_{efgh} \Rightarrow & \sum (-1)^x \gamma^{ab}_{ef} \gamma^{cd}_{gh} (18 \text{ terms}) - 2 \sum (-1)^x \gamma^a_e \gamma^b_f \gamma^c_g \gamma^d_h (24 \text{ terms}). \qquad (5.11)
\end{aligned}$$

Note that in the case of the four-particle density matrix we have not written out all terms explicitly, but these can be obtained by attaching appropriate signs to the additional contributions arising from permutation of indices: this is denoted by the parity factor $(-1)^x$ which takes the appropriate $+1, -1$ sign according to the permutation. To obtain the cu-SC-NEVPT2 energy and coefficients, we simply substitute the approximate 3- and 4-particle density matrices as defined above into the matrix element expressions in NEVPT2 (e.g. as defined in the appendix of Ref. [29]). Note that the contributions of subspaces $(0), (+1), (-1)$ are not affected by the cumulant approximation.

cud-SC-NEVPT2: To improve on the cumulant approximated strongly-contracted NEVPT2, we have investigated a better approximation to the 3- and 4-particle density matrices. In a typical basis, the largest elements of the 3-particle and 4-particle density matrices occur along the diagonals and thus it would appear most important to include information about these elements first. This can be done by incorporating the exact 3- and 4-particle diagonal elements of the density matrices. This does not increase the complexity of the theory over that of cu-SC-NEVPT2 since the number of 4-particle diagonal elements

(from which the 3-particle diagonals are readily obtained) is n_{act}^4, which is the same number of elements as in the two-particle density matrix. We can also incorporate some additional information of the same complexity concerning the off-diagonal elements of the 3- and 4-particle density matrices, as long as we consider subsets of elements which are labeled by 4 indices or fewer. Thus, in the cumulant with diagonals approximated SC-NEVPT2 theory, we construct the 3-particle density matrix via the cumulant approximation Eqns. 5.10 and 5.11 but then replace selected diagonal and off-diagonal terms with their exact values, in particular elements $\gamma_{abc}^{abc}, \gamma_{bac}^{abc}, \gamma_{cab}^{abc} \ldots$. In the case of the 4-particle density matrix, since we now have explicit information from the 3-particle density matrix beyond the cumulant approximation, to be consistent we construct the 4-particle density matrix from its cumulant expansion in terms of 1, 2, and 3-particle density matrices where the 3-particle density matrix has been corrected as above. This expansion is given by

$$\begin{aligned}
\gamma_{efgh}^{abcd} = &\sum (-1)^x \gamma_e^a \gamma_{fgh}^{bcd} (16 \text{ terms}) \\
&+ \sum (-1)^x \gamma_{ef}^{ab} \gamma_{gh}^{cd} (18 \text{ terms}) \\
&- 2 \sum (-1)^x \gamma_e^a \gamma_f^b \gamma_{gh}^{cd} (72 \text{ terms}) \\
&+ 6 \sum (-1)^x \gamma_e^a \gamma_f^b \gamma_g^c \gamma_h^d (24 \text{ terms}).
\end{aligned} \tag{5.12}$$

From this cumulant approximated 4-particle density matrix, we then replace the following four-indexed diagonal and off-diagonal elements with their exact values, $\gamma_{abcd}^{abcd}, \gamma_{bacd}^{abcd}, \gamma_{cabd}^{abcd} \ldots$.

In our implementation the 4-particle density matrix is not kept in memory due to earlier precontraction with integrals of the 1-, 2-, and 3- particle density matrices used in the cumulant expansion, which lowers the overall computational scaling.

5.3.3 False intruders from cumulant approximations

Intruder states are perturber states (i.e. excited from the reference) which are near degenerate with the reference configuration with respect to the zeroth order Hamiltonian (but are not so with the true Hamiltonian). They give rise to a zeroth order problem which appears near-degenerate when no such near-degeneracy exists in the real system and are a common problem in multireference perturbation theories such as CASPT2. NEVPT2 theory was developed in part to address the problem of intruder states. The use of the two-electron Hamiltonian in the active space (via the zeroth order Dyall Hamiltonian) ensures a better estimate of the relative energies of the perturber states and reference state, reducing the risk of possible intruders.

However, in the cumulant based approximations that we have introduced, intruder states can re-appear. This is because the energy denominators in the NEVPT2 theory are no longer evaluated exactly, and the errors of the cumulant approximations may lead to falsely small denominators. We refer to intruders that arise in this way as "false intruders". Consider, for example, the expression for the second-order energy in the $(-1)'$ subspace. Rewriting Eqn. (5.9) using the fact that $H_{\text{act}}\Psi^{[0]} = E^{[0]}\Psi^{[0]}$, we have

$$\begin{aligned} E^{[2]} &= -\sum_r \frac{N_r}{E_r - E^{[0]}} \\ &= -\sum_r \frac{N_r}{\epsilon_r - \langle \Psi^{[0]} | V_r^{(-1)'\dagger} [H_D, V_r^{(-1)'}] | \Psi^{[0]} \rangle}. \end{aligned} \tag{5.13}$$

Both the numerator and the denominator are approximated, the numerator requiring a 3-particle density matrix and the denominator (by virtue of the $\langle \Psi^{[0]} | V_r^{(-1)'\dagger} [H_D, V_r^{(-1)'}] | \Psi^{[0]} \rangle$ term) requiring a four-particle density matrix. We can imagine two kinds of errors introduced by the cumulant approximation to

the denominator. In the first case, the denominators may all simply be poor e.g. somewhat shifted, but no divergences occur. In the second case, the the cumulant approximation to $\langle \Psi^{[0]} | V_r^{(-1)'\dagger} [H_D, V_r^{(-1)'}] | \Psi^{[0]} \rangle \approx \epsilon_r$ and a false intruder appears, with a corresponding divergent contribution to the energy.

To ameliorate the effects of possible false intruders, we have incorporated level shifts into our cu-SC-NEVPT2 and cud-SC-NEVPT2 methods. We use imaginary level shifts as investigated by Forsberg and Malmqvist [31]. When using the level shift, we have also evaluated the cumulant approximated NEVPT2 energies using the level-shift corrected energy expression of Ref. [31], which is designed to minimize the effect of the level-shift when there are no intruder states. The correction corresponds to evaluating the second-order energy from the Hylleraas functional with the first-order wavefunction coefficients determined with the level-shift. Note that when using level-shifts with cumulant approximated SC-NEVPT2 we need not apply the level-shift to every subspace, only the subspace in which the false intruder is observed. In the tables and figures, we list the value of the level-shift in parentheses e.g. cu-SC-NEVPT2(0.2) means the value of the applied imaginary level shift was $0.2i\ a.u.$.

5.4 Test-cases

We have incorporated the cu-SC-NEVPT2 and cud-SC-NEVPT2 approximations into the existing NEVPT2 implementation in the development version of DALTON [32]. To assess the accuracy of these approximations, we have studied a number of benchmark quantum chemistry problems: the singlet-triplet gap of CH_2 and SiH_2, the nitrogen and chromium dimer potential energy curves, and

excitation energies in short polyenes. Unless otherwise specified, benchmark results for ordinary SC-NEVPT2 were obtained with DALTON, while benchmark results for CASSCF, FCI, and CASPT2 were obtained with MOLPRO [33]. (In the latter case the "rs2" variant was used in all cases except for the polyenes, where the "rs2c" variant" was used).

5.4.1 Test-case I: Singlet-triplet gaps in CH_2 and SiH_2

As a first test of the accuracy of the cumulant approximated SC-NEVPT2 methods, we calculated the ground-state singlet-triplet splittings in CH_2 and SiH_2. These are small quantities and thus very sensitive to any errors made in the differential correlation between the singlet and triplet states. The same systems have been used in earlier studies to benchmark the accuracy of multireference perturbation theory [20].

For CH_2, the geometry was taken from Ref. [34] and the basis set (double-zeta with polarization quality) was also taken from Ref. [34]. (Note that different polarization functions are used for the singlet and triplet states for CH_2). The CASSCF active space consisted of $2a_1, 3a_1, 4a_1$, $1b_1, 2b_1$, and $1b_2$ orbitals. The $1a_1$ orbital was a core orbital (always doubly occupied) in the CASSCF calculation and was treated as a frozen (i.e. an uncorrelated orbital) in all NEVPT2, CASPT2, and FCI calculations. The energies of the $^1A_1$ and ^3B_u states were obtained at the state-specific CASSCF level, and these states were subsequently used in the NEVPT2 and CASPT2 calculations.

For SiH_2, the geometry was taken from Ref. [35] and the basis set (which was of double-zeta with polarization quality) was also taken from Ref. [35]. The

Table 5.2: Ground-state singlet-triplet gap in CH_2 (see text for basis and geometry). The values in brackets with cu-SC-NEVPT2 and cud-SC-NEVPT2 denote the imaginary level shift in units of $a.u.$ used.

Method	Singlet energy/E_h	Triplet energy/E_h	Gap kcal/mol
FCI	-39.027183	-39.046229	11.95
CASSCF	-38.945529	-38.965954	12.82
CASPT2	-39.012184	-39.037061	15.61
SC-NEVPT2	-39.006707	-39.028498	13.67
cu-SC-NEVPT2(0.1)	-39.003499	-39.042508	24.48
cud-SC-NEVPT2(0.1)	-39.006470	-39.029253	14.30

CASSCF active space consisted of $4a_1, 5a_1, 6a_1$, $2b_1, 3b_1$, $2b_2$ orbitals, with all lower orbitals held as core orbitals in the CASSCF calculation, and frozen in subsequent NEVPT2, CASPT2, and FCI calculations. The energies of the $^1A_1$ and ^3B_u states were obtained at the state-specific CASSCF level, and these states were subsequently used in the NEVPT2 and CASPT2 calculations.

Discussion: As can be seen from tables 5.2, 5.3, the pure cumulant based cu-SC-NEVPT2 result is poorly balanced between the singlet and triplet states, while the cumulant with diagonals cud-SC-NEVPT2 approximation performs significantly better. In fact, the cu-SC-NEVPT2 gaps are worse than the CASSCF gaps! The cud-SC-NEVPT2 gaps are of similar quality (relative to the FCI result) to the standard SC-NEVPT2 or CASPT2 singlet-triplet gaps.

Table 5.3: Ground-state singlet-triplet gap in SiH_2 (see text for basis and geometry). The values in brackets with cu-SC-NEVPT2 and cud-SC-NEVPT2 denote the imaginary level shift in units of $a.u.$ used.

Method	Singlet energy / E_h	Triplet energy / E_h	Gap kcal/mol
FCI	-290.110206	-290.082219	17.50
CASSCF	-290.042910	-290.016811	16.38
CASPT2	-290.095403	-290.071351	15.09
SC-NEVPT2	-290.088824	-290.062222	16.69
cu-SC-NEVPT2(0.1)	-290.083694	-290.072706	6.90
cud-SC-NEVPT2(0.1)	-290.089888	-290.064103	16.18

To better understand the origin of the errors in the cumulant approximated methods, we can analyze the contributions to the errors from each subspace of the NEVPT2 calculations, which we do now for the CH_2 singlet and triplet states. Since all the core orbitals are frozen (uncorrelated) in the CH_2 calculation, all correlation energies in subspaces that involve excitations to and from core orbitals vanish. This means that only subspaces (-2) and $(-1)'$ contribute to the correlation energy. We compare the first-order wavefunction norm and energy contributions of these subspaces for the singlet and triplet states and the various SC-NEVPT2 approximations in Table 5.4.

We see that in both cu-SC-NEVPT2 and cud-SC-NEVPT2, the first-order wavefunction norm and second-order energy contribution of the (-2) space is well approximated; errors in the correlation energy for this space are less than 0.5 mH. From table 5.1 we recall that the (-2) subspace requires a 3-particle cumulant approximation in the denominators appearing in the NEVPT2 coefficient

Table 5.4: Subspace contributions to the correlation energy and wave-function norm for various SC-NEVPT2 methods in the CH_2 molecule.

State	Subspace	SC-NEVPT2		cu-SC-NEVPT2		cud-SC-NEVPT2	
		Wf norm	Energy/mE_h	Wf norm	Energy/mE_h	Wf norm	Energy/mE_h
$^1A_1$	(-2)	7.5×10^{-3}	-31.8	8.4×10^{-3}	-31.3	7.4×10^{-3}	-31.7
	$(-1)'$	11.9×10^{-3}	-29.4	16.2×10^{-3}	-26.7	11.9×10^{-3}	-29.2
$^3B_1$	(-2)	6.9×10^{-3}	-30.9	6.9×10^{-3}	-30.9	6.9×10^{-3}	-30.9
	$(-1)'$	12.1×10^{-3}	-31.6	40.0×10^{-3}	-45.6	12.9×10^{-3}	-32.4

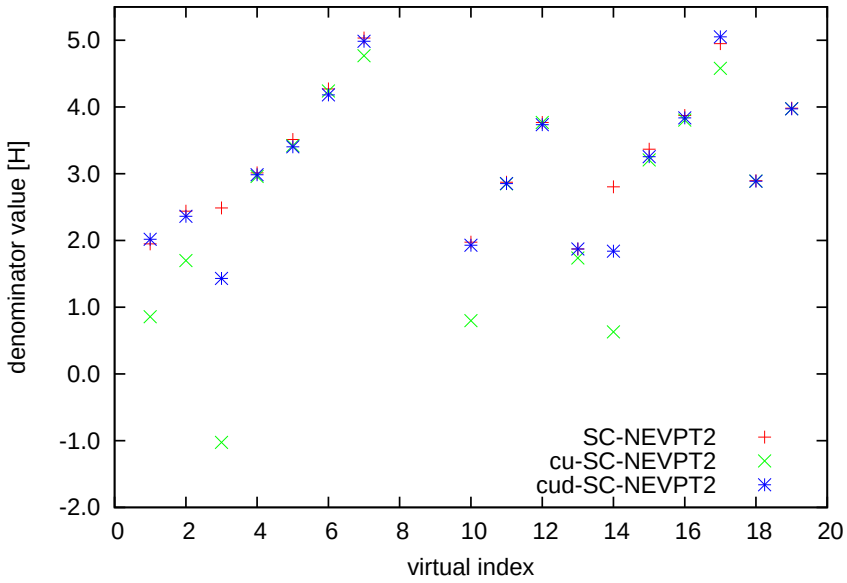


Figure 5.1: Energy denominators for respective virtual indices in the $(-1)'$ subspace for the $^3B_1$ state of CH_2 using SC-NEVPT2, cu-SC-NEVPT2, cud-SC-NEVPT2.

and energy expressions. The largest errors arise in the $(-1)'$ space. In particular, the cu-SC-NEVPT2 correlation energy in this space for the triplet state is over-estimated by -14 mE_h, giving rise to the spuriously large singlet-triplet gap. Consistent with this is the error in the wavefunction norm in this subspace, which is more than twice the correct norm. From table 5.1 we recall that the $(-1)'$ subspace requires *both* 3- and 4-particle cumulant approximations, and is thus expected to be associated with larger errors.

We can further break up the error contribution by the error in the numerator and denominator of the sum-over-states expression for the energy Eqn. 5.13. Although the numerators that require at most the 3-particle density matrix are approximated well using the cumulant approximation, the main source of error lies in the values of the denominators that are usually too small. In Fig. 5.1

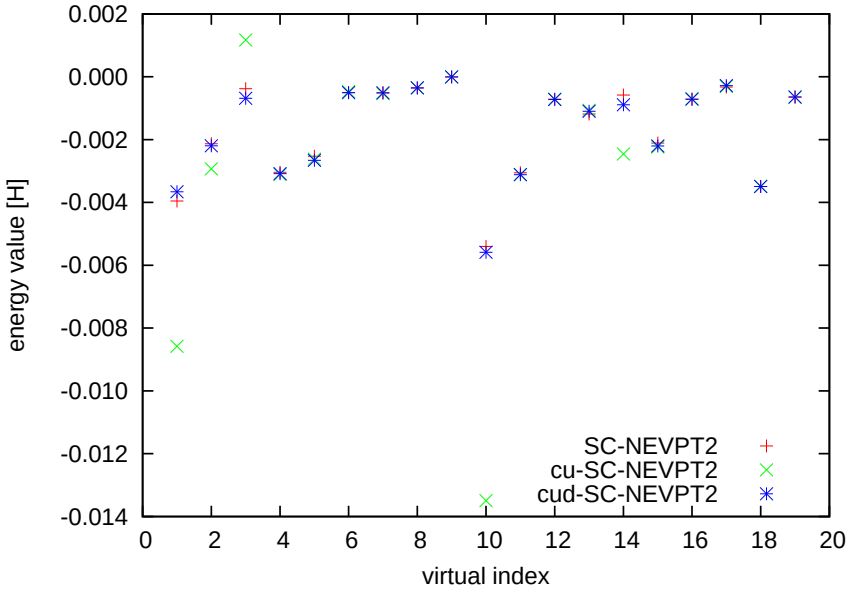


Figure 5.2: Energy contributions for respective virtual indices in the $(-1)'$ subspace for the $^3B_1$ state of CH_2 using SC-NEVPT2, cu-SC-NEVPT2, cud-SC-NEVPT2.

we plot (as a function of virtual index) the denominators for the $(-1)'$ subspace in the SC-NEVPT2, cu-SC-NEVPT2 and cud-SC-NEVPT2 theories. While the cud-SC-NEVPT2 denominators appear to be close to the SC-NEVPT2 ones, the denominators obtained without introduction of the diagonal elements (i.e. for cu-SC-NEVPT2) seem to be too small for many of the virtual indices. However, although we see that errors in the denominators of the cu-SC-NEVPT2 expressions no doubt contribute to the overly large wavefunction norm and overestimated correlation energies in this subspace, we do not really have a case of a vanishing denominator and diverging energy (see Fig. 5.2) contribution due to a false intruder state. For this reason, the effect of a level-shift correction is quite small. Indeed, the cu-SC-NEVPT2 singlet-triplet gap calculated with imaginary level shifts of 0.1, 0.2, and 0.4 $a.u.$ in the $(-1)'$ subspace, calculated with the corrected level-shift energies as in Ref. [31], remains 24.48 kcal/mol. One can

see that a better estimation of the denominator is needed than that given by the pure cumulant approximation. It appears, that for these systems, the primary errors in the cumulant approximation can be corrected by incorporating the additional "diagonal" elements, as used in the cud-SC-NEVPT2 method, and that this method gives a qualitatively balanced description of the dynamic correlation, but the results may be less satisfactory in other cases.

5.4.2 Test-case II: Nitrogen and chromium potential energy curves

As a more stringent evaluation of the cu-SC-NEVPT2 and cud-SC-NEVPT2 approximations, we carried out calculations on the nitrogen and chromium binding curves. The correct description of multiple bond-stretching is amongst the hardest problems in benchmark quantum chemistry. Furthermore, it is well known from CASPT2 studies that intruder state problems can arise at many geometries along such potential energy curves [36].

For N_2 we studied the lowest singlet $^1\Sigma_g$ and lowest triplet $^3\Sigma_u$ states. We used Dunning's correlation consistent quadruple zeta basis (cc-pVQZ) [37, 38] and the CASSCF active space consisted of all $2s$ and $2p$ orbitals (10e, 8o) active space. The $1s$-derived orbitals were kept doubly occupied in the CASSCF reference and correlated in the subsequent CASPT2 and NEVPT2 calculations as core orbitals. For Cr_2 we studied the lowest singlet state using the Wachters+f atomic natural orbital basis [37, 39, 40] and a CASSCF active space containing all $3d$ and $4s$ orbitals (12e, 12o) active space. All other occupied orbitals were correlated as core orbitals.

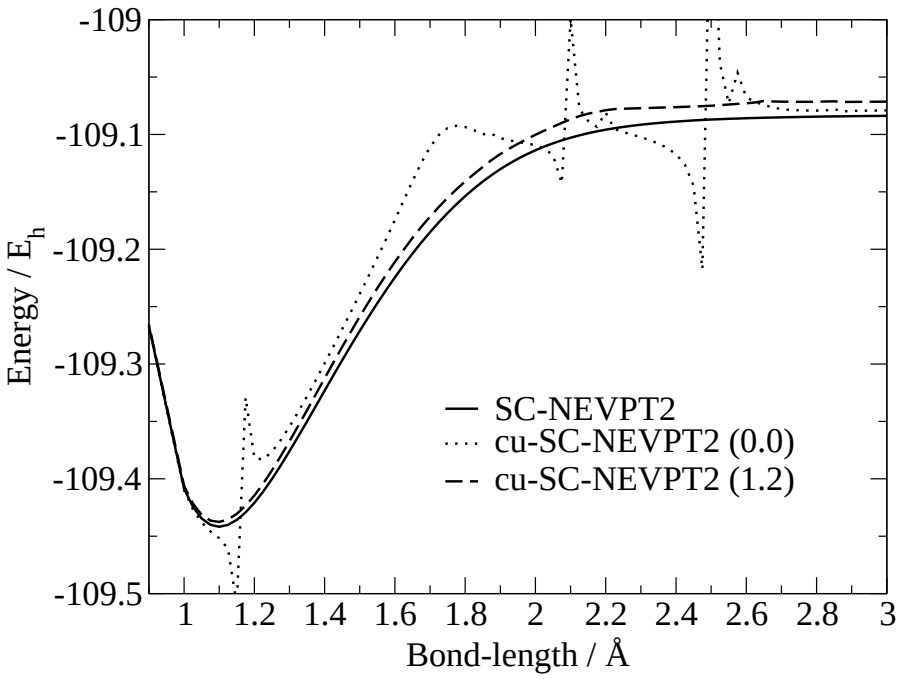


Figure 5.3: ${}^1\Sigma_g$ binding curve of N_2 using cu-SC-NEVPT2 and NEVPT2 theory. The cu-SC-NEVPT2 calculations are carried out using a variety of imaginary level shifts in the $(-1)'$ subspace (value indicated in brackets).

Discussion: Shown in Fig. 5.3 is the ${}^1\Sigma_g$ binding curve using both cu-SC-NEVPT2 and SC-NEVPT2 theory. What is immediately apparent is that the cu-SC-NEVPT2 curve has several divergences arising from false intruder states which do not appear in the original SC-NEVPT2 theory. These divergences occur even close to the equilibrium region, but can be smoothed out by applying a sufficiently large imaginary level shift in the $(-1)'$ subspace. The magnitude of the necessary level shift (1.2 $a.u.$) may seem rather large as compared to typical shifts employed in CASPT2 theory. However, the physical meaning of the contribution $\langle \Psi^{[0]} | V_r^{(-1)'\dagger} [H_D, V_r^{(-1)'}] | \Psi^{[0]} \rangle$ in the $(-1)'$ subspace is that of a generalized ionization energy, and, as observed in Fig. 5.1, the typical value of such energies in the parent SC-NEVPT2 theory is $O(1)$ $a.u.$. Thus it is natural

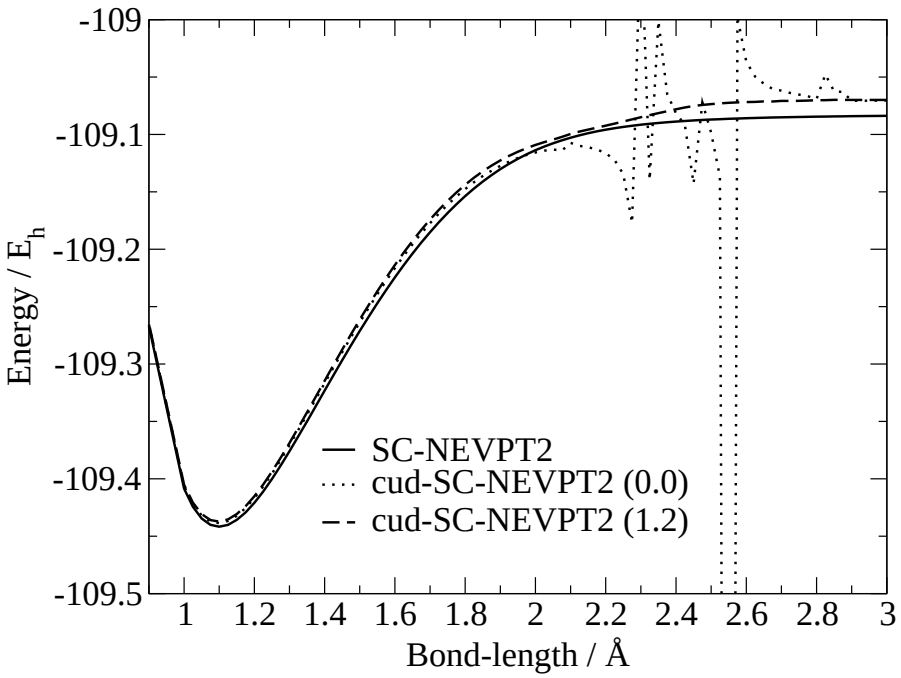


Figure 5.4: ${}^1\Sigma_g$ binding curve of N_2 using cud-SC-NEVPT2 and NEVPT2 theory. The cu-SC-NEVPT2 calculations are carried out using a variety of imaginary level shifts in the $(-1)'$ subspace (value indicated in brackets).

to require a level shift of this magnitude to correct divergences. Such a large level shift however does give rise to a significantly vertically shifted curve in the dissociation region.

In Fig. 5.4 is the corresponding ${}^1\Sigma_g$ binding curve using cud-SC-NEVPT2 and SC-NEVPT2 theory. The cud-SC-NEVPT2 curve does not suffer from false intruders in the equilibrium region and instead closely follows the parent SC-NEVPT2 curve. However, at approximately 2.2Å we observe divergences once again. These can be also be smoothed out using an imaginary level shift, leading to a similarly vertically shifted curve in the dissociation region. In Fig. 5.5 we plot the ${}^3\Sigma_u$ binding curve using cud-SC-NEVPT2 and SC-NEVPT2 theory. Note that unlike CASPT2 calculations in ANO basis sets [36], SC-NEVPT2 does not

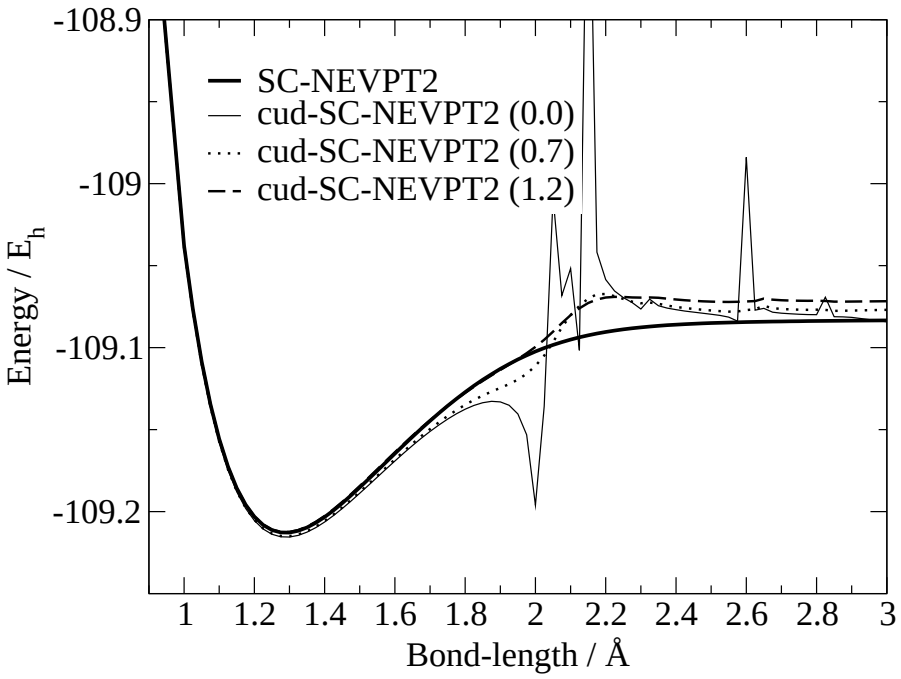


Figure 5.5: ${}^3\Sigma_u$ binding curve of N_2 using cud-SC-NEVPT2 and NEVPT2 theory. The cud-SC-NEVPT2 calculations are carried out using a variety of imaginary level shifts in the $(-1)'$ subspace (value indicated in brackets).

exhibit any intruder problems for this state. As for the ground state, there are no false intruders in the cud-SC-NEVPT2 curve until one reaches the stretched region, near 2Å. Although it is possible to remove these intruders with a large imaginary level shift, the resulting smoothed curve has a very different shape in the stretched region and is not entirely satisfactory. It should not be surprising that the quality of the cud-SC-NEVPT2 curve degrades at longer distances. At such stretched geometries, there are many fractional occupation numbers and consequently, the 3-particle and 4-particle density matrices are less diagonally dominant, even in the natural orbital basis. Thus, the amount of information provided by incorporating the diagonal elements of these quantities is decreased.

In Table 5.5 we show the spectroscopic constants (obtained by numerical fitting of the potential energy curve) for the $^1\Sigma_g$ and $^3\Sigma_u$ states of N_2. While the bare (i.e. without level-shift) cu-SC-NEVPT2 spectroscopic constants are quite poor, application of the imaginary level shift, which smoothes over the equilibrium region of the curve actually produces quite reasonable spectroscopic constants of an accuracy comparable to standard CASPT2 and SC-NEVPT2. Only the dissociation energy is somewhat large, and this can be attributed to the vertical shift of the dissociated region of the curve due to the level shift as discussed above. cud-SC-NEVPT2 spectroscopic constants with or without level shifts are quite reasonable for both states, and are again comparable in accuracy to standard CASPT2 and SC-NEVPT2. This reflects the relatively good behavior of the cud-SC-NEVPT2 approximation in the equilibrium region.

In Fig. 5.6 we show the Cr_2 binding curve computed using SC-NEVPT2 and the cud-SC-NEVPT2 approximation with a variety of imaginary level shifts. In this challenging system, even the cud-SC-NEVPT2 approximation shows strong false intruder behavior in the equilibrium region. By applying successively larger imaginary level shifts, we can smooth out the divergences, but the general quality of the potential energy curve is not so good. At larger level shifts, we obtain a double well with cud-SC-NEVPT2, rather than the shouldered single-well type curve that is believed to characterize Cr_2, and in particular the well at longer bond-distances appears deeper than the well at the normal Cr_2 bond-length. It should be noted, however, that the Cr_2 binding curve is very sensitive to the level of theory employed, and, for example, SC-NEVPT3 [41], CIPT2 [42], and internally contracted CI using the n-electron valence states [41] all produce curves with a double well structure not unlike our approximate cud-SC-NEVPT2 curve.

Table 5.5: Spectroscopic constants for the nitrogen molecule calculated with CASSCF, CASPT2, SC-NEVPT2, cu-SC-NEVPT2 and cud-SC-NEVPT2. The imaginary level shifts (in $a.u.$) given in brackets.

$N_2\ ^1\Sigma_g$			
Method	r_e/Å	D_e/eV	$\omega_e/\mathrm{cm}^{-1}$
Expt.	1.0977	9.91	2359
CASSCF	1.1069	9.23	2496
CASPT2	1.1012	9.51	2454
SC-NEVPT2	1.1021	9.77	2460
cu-SC-NEVPT2(0.0)	1.1537	11.24	3930
cu-SC-NEVPT2(1.2)	1.0980	9.94	2470
cud-SC-NEVPT2(0.0)	1.1002	9.87	2466
cud-SC-NEVPT2(1.2)	1.0997	9.98	2467
$N_2\ ^3\Sigma_u$			
Method	r_e/Å	D_e/eV	$\omega_e/\mathrm{cm}^{-1}$
Expt.	1.2866	3.68	1461
CASSCF	1.3027	2.79	1548
CASPT2	1.2879	3.56	1513
SC-NEVPT2	1.2905	3.54	1522
cud-SC-NEVPT2(0.0)	1.2922	3.78	1521
cud-SC-NEVPT2(0.7)	1.2917	3.79	1522
cud-SC-NEVPT2(1.2)	1.2892	3.82	1524

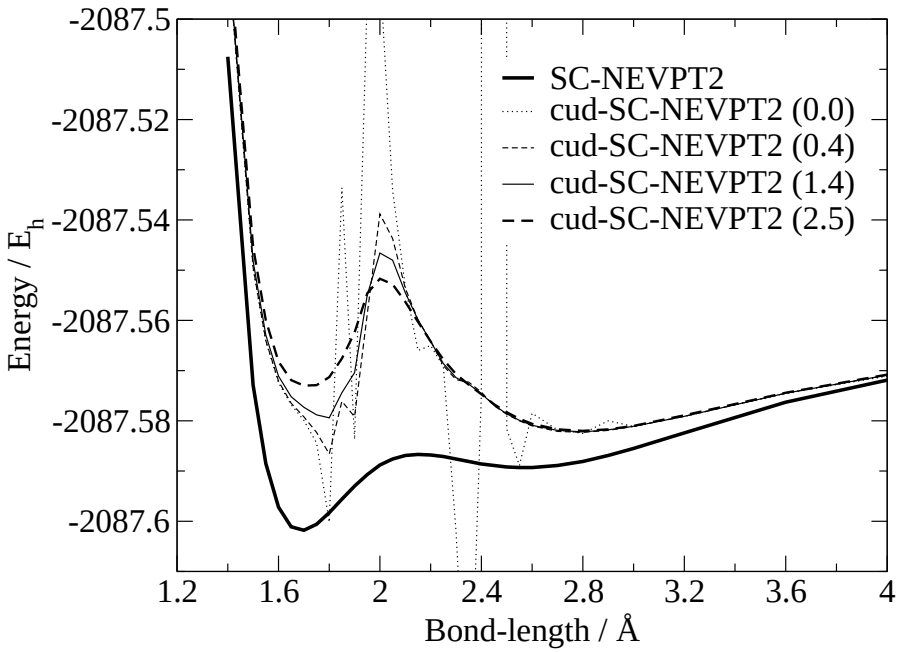


Figure 5.6: Cr_2 binding curve using cud-SC-NEVPT2 and SC-NEVPT2 theory. The cud-SC-NEVPT2 calculations are carried out using a variety of imaginary level shifts in the $(-1)'$ subspace (value indicated in brackets).

5.5 Test-case III: Excited states in polyenes

An area of considerable success for multireference perturbation theory calculations, particularly those employing CASPT2 or MRMP theory, has been the description of excited states of small to medium sized organic molecules [43, 44]. As examples of such systems, here we have chosen short-chain polyenes to assess the behavior of the cumulant approximations to SC-NEVPT2.

The geometries of the C_4H_6, C_6H_8, C_8H_{10}, $C_{10}H_{12}$ in the all-trans configurations were optimized at the DFT/B3LYP [45, 46] using Dunning's correlation consistent cc-pVDZ basis [37, 38]. The subsequent wavefunction calculations were carried out in Dunning's correlation consistent cc-pVDZ basis. The CASSCF ac-

tive space was chosen to be the full π-valence space. The excited CASSCF states as listed in tables 5.6, 5.7, 5.8 and 5.9 for the NEVPT2 calculations were obtained by state-specific CASSCF using the DALTON CASSCF algorithm, while the excited CASSCF states for the CASPT2 comparison calculations were obtained through state-averaged CASSCF using the MOLPRO CASSCF algorithm. For the state-averaged CASSCF, the state-average incorporated the lowest 3 states of A_g symmetry (for the A_g calculations) and the lowest 2 states of B_u symmetry (for the B_u calculations) while the CASPT2 correction was calculated in a state-specific way (i.e. the density matrix of the given state, rather than the average density matrix was used in the construction of the zeroth order Hamiltonian). The CASPT2 calculations used the "rs2c" variant as implemented in MOLPRO. In all CASPT2 and NEVPT2 calculations the σ electrons were correlated as core orbitals.

Discussion: The primary effect of dynamical correlation on the low-lying valence excited states in polyenes is to lower the energy of the "ionic" excited state $1B_u^+$ relative to the covalent excited states $2A_g^-$, $3A_g^-$, $1B_u^-$. Of particular interest is the crossing point, i.e. the length of polyene at which the $1B_u^+$ state becomes degenerate with the $2A_g^-$ state. Comparing the cu-SC-NEVPT2, cud-SC-NEVPT2, and parent SC-NEVPT2 methods in tables 5.6, 5.7, 5.8 and 5.9 we observe that the cumulant approximated methods do reproduce the lowering of the $1B_u^+$ state relative to the covalent excited states (as compared to CASSCF). Furthermore, the cu-SC-NEVPT2 and cud-SC-NEVPT2 preserve the state-ordering of the parent SC-NEVPT2 method, and with these methods the $1B_u^+$ and $2A_g^-$ states become nearly degenerate in $C_{10}H_{12}$ just as in the parent SC-NEVPT2 method. The cumulant approximated excitation energies are consistently too low compared to the SC-NEVPT2 excitation energies, and this un-

Table 5.6: Low-lying valence excitations in butadiene with CASSCF, CASPT2, SC-NEVPT2, cu-SC-NEVPT2 and cud-SC-NEVPT2. Basis and geometry described in text.

Molecule	Method	Energy / E_h	Excitations / eV			
		$1A_g^-$	$2A_g^-$	$3A_g^-$	$1B_u^-$	$1B_u^+$
C_4H_6	Expt.					6.25 [1], 5.92 [2]
	CASSCF	-154.98880	6.75	10.93	13.20	8.46
	CASPT2	-155.47683	6.49	8.14	10.84	6.27
	SC-NEVPT2	-155.48615	6.91	9.25	11.30	6.46
	cu-SC-NEVPT2	-155.48333	6.75	8.86	11.19	6.20
	cud-SC-NEVPT2	-155.48230	6.80	8.66	11.41	6.35

[a] [47]
[b] [48–50]

Table 5.7: Low-lying valence excitations in hexatriene with CASSCF, CASPT2, SC-NEVPT2, cu-SC-NEVPT2 and cud-SC-NEVPT2. Basis and geometry described in text.

Molecule	Method	Energy / E_h	Excitations / eV			
		$1A_g^-$	$2A_g^-$	$3A_g^-$	$1B_u^-$	$1B_u^+$
C_6H_8	Expt.					4.93 [1], 4.95 [2], 5.13 [3]
	CASSCF	-231.91053	5.61	8.79	7.37	6.77
	CASPT2	-232.63634	5.21	8.35	5.08	6.32
	SC-NEVPT2	-232.64938	5.60	9.00	5.35	6.80
	cu-SC-NEVPT2	-232.64434	5.38	8.75	5.12	6.65
	cud-SC-NEVPT2	-232.64134	5.37	8.81	5.13	6.60

[a] [51, 52]
[b] [51, 52]
[c] [53, 54]

Table 5.8: Low-lying valence excitations in octatetraene with CASSCF, CASPT2, SC-NEVPT2, cu-SC-NEVPT2 and cud-SC-NEVPT2. Basis and geometry described in text.

Molecule	Method	Energy / E_h	Excitations / eV			
		$1A_g^-$	$2A_g^-$	$3A_g^-$	$1B_u^-$	$1B_u^+$
C_8H_{10}	Expt.		3.54 [1]			4.41 [2]
	CASSCF	-308.83237	4.83	6.73	6.03	6.70
	CASPT2	-309.79536	4.35	6.05	5.47	4.40
	SC-NEVPT2	-309.81313	4.71	6.67	5.94	4.47
	cu-SC-NEVPT2	-309.80488	4.44	6.42	5.68	3.69
	cud-SC-NEVPT2	-309.79906	4.30	6.30	5.56	4.04

[a] [55]
[b] [56, 57]

Table 5.9: Low-lying valence excitations in decapentaene with CASSCF, CASPT2, SC-NEVPT2, cu-SC-NEVPT2 and cud-SC-NEVPT2. Basis and geometry described in text.

Molecule	Method	Energy / E_h	Excitations / eV			
		$1A_g^-$	$2A_g^-$	$3A_g^-$	$1B_u^-$	$1B_u^+$
$C_{10}H_{12}$	Expt.		3.48 [1]			4.02 [2]
	CASSCF	-385.75357	4.20	6.11	5.27	6.22
	CASPT2	-386.95568	3.68	5.39	4.70	3.91
	SC-NEVPT2	-386.97763	4.03	6.01	5.15	3.90
	cu-SC-NEVPT2	-386.96667	3.59	5.66	4.86	3.52
	cud-SC-NEVPT2	-386.95549	3.40	5.39	4.42	3.28

[a] [58]
[b] [58]

derestimation appears to get worse as the polyene chain gets longer. Most of the excitation error can be traced to the error in the ground-state energy: both cu-SC-NEVPT2 and cud-SC-NEVPT2 place the energy of the ground-state too high, leading to an overall decrease in all the excitation energies. Examining the contributions of the different subspaces, once again we observe that the largest error in the cumulant approximated methods occurs in the $(-1)'$ subspace. Unlike in the previous test-cases, the cud-SC-NEVPT2 method does not perform any better than the cu-SC-NEVPT2 method. Overall, we observe that the cumulant approximated theories give a qualitatively reasonable picture of the excitation energies in these conjugated molecules, although the quantitative accuracy for large conjugated systems remains to be seen.

5.6 Conclusions

In this chapter, we have explored the possibility of constructing approximations to multireference perturbation theory that do not depend on three- and four-particle density matrices, with the view to enabling dynamical correlation calculations in conjunction with very large active spaces. As our parent multireference perturbation theory, we have investigated the strongly-contracted variant of the n-electron valence perturbation theory. Our strategy has been to employ cumulant type approximations to the three- and four-particle density matrices that appear in the formulation using the one- and two-particle density matrices and quantities of similar complexity. We have proposed two cumulant approximated methods: cumulant approximated SC-NEVPT2 (cu-SC-NEVPT2) and cumulant with diagonals approximated SC-NEVPT2 (cud-SC-NEVPT2). The latter incorporates additional exact information about diagonal and off-diagonal

elements of the three- and four-particle density matrices with the same complexity $O(n_{act}^4)$ as the two-particle density matrices. We find that an undesirable feature introduced by using cumulant approximations is the re-emergence of intruder states in the perturbation theory (which do not usually appear in the strongly-contracted NEVPT2 theory) due to the inaccurate representation of denominators by their cumulant approximated form. We have assessed the cumulant approximations in several benchmark test systems. We find that the cumulant approximated methods, when augmented by appropriate level shifts to deal with possible intruder state problems do provide a qualitatively correct picture of dynamical correlation in many cases. We find also that the cud-SC-NEVPT2 theory has much weaker intruder state problems than the pure cumulant cu-SC-NEVPT2 theory. The accuracy of the cumulant approximated theories is necessarily degraded from the parent multireference perturbation theory, although in many cases the cumulant derived error is within the intrinsic error range associated with multireference second-order perturbation theory. While the cumulant approximated theories may be used with care as a practical means to obtain information on qualitative effects of dynamical correlation in systems with many active orbitals where the parent multireference perturbation theories cannot be applied, we would clearly like more reliable approximations to the denominators in the SC-NEVPT2 method.

BIBLIOGRAPHY

[1] D. Zgid, D. Ghosh, E. Neuscamman, and G. K.-L. Chan, J. Chem. Phys. **130**, 194107 (2009).

[2] L. Meissner, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **91**, 6187 (1989).

[3] N. Oliphant and L. Adamowicz, J. Chem. Phys. **94**, 1229 (1991).

[4] U. S. Mahapatra, B. Datta, and D. Mukherjee, J. Chem. Phys. **110**, 6171 (1999).

[5] K. Andersson, P.-Å. Malmqvist, and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).

[6] Y.-K. Choe, Y. Nakao, and K. Hirao, J. Chem. Phys. **115**, 621 (2001).

[7] K. Hirao, Chem. Phys. Lett. **190**, 374 (1992).

[8] C. Angeli, R. Cimiriglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, J. Chem. Phys. **114**, 10252 (2001).

[9] S. R. White and R. L. Martin, J. Chem. Phys. **110**, 4127 (1999).

[10] K. H. Marti, I. M. Ondik, G. Moritz, and M. Reiher, J. Chem. Phys. **128**, 014104 (2008).

[11] D. Zgid and M. Nooijen, J. Chem. Phys. **128**, 144115 (2008).

[12] D. Zgid and M. Nooijen, J. Chem. Phys. **128**, 144116 (2008).

[13] D. Ghosh, J. Hachmann, T. Yanai, and G. K.-L. Chan, J. Chem. Phys. **128**, 144117 (2008).

[14] J. Hachmann, J. J. Dorando, M. Aviles, and G. K.-L. Chan, J. Chem. Phys. **127**, 134309 (2007).

[15] J. Hachmann, W. Cardoen, and G. K.-L. Chan, J. Chem. Phys. **125**, 144101 (2006).

[16] G. K.-L. Chan, J. Chem. Phys. **120**, 3172 (2004).

[17] P.-Å. Malmqvist, A. Rendell, and B. O. Roos, J. Phys. Chem. **94**, 5477 (1990).

[18] D. Walter and E. A. Carter, Chem. Phys. Lett. **346**, 177 (2001).

[19] P. Celani and H. J. Werner, J. Chem. Phys. **112**, 5546 (2000).

[20] R. W. A. Havenith, P. R. Taylor, C. Angeli, R. Cimiraglia, and K. Ruud, J. Chem. Phys. **120**, 4619 (2004).

[21] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys **107**, 432 (1997).

[22] F. Colmenero and C. Valdemoro, Phys. Rev. A **47**, 979 (1993).

[23] H. Nakatsuji and K. Yasuda, Phys. Rev. Lett. **76**, 1039 (1996).

[24] D. A. Mazziotti, Phys. Rev. A **57**, 4219 (1998).

[25] D. A. Mazziotti, editor, *Reduced-Density-Matrix Mechanics with Application to Many-electron Atoms and Molecules*, volume 134 of *Adv. Chem. Phys.*, Wiley, New York, 2007.

[26] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **127**, 104107 (2007).

[27] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **124**, 194106 (2006).

[28] S. R. White, J. Chem. Phys. **117**, 7472 (2002).

[29] C. Angeli, R. Cimiriglia, and J.-P. Malrieu, J. Chem. Phys. **117**, 9138 (2002).

[30] K. G. Dyall, J. Chem. Phys. **102**, 4909 (1995).

[31] N. Forsberg and P.-Å. Malmqvist, Chem. Phys. Lett. **274**, 196 (1997).

[32] Dalton, a molecular electronic structure program, release 2.0 (2005), see http://www.kjemi.uio.no/software/dalton/dalton.html.

[33] H.-J. Werner et al., Molpro, version 2008.1, a package of ab initio programs, 2008, see http://www.molpro.net.

[34] C. W. Bauschlicher Jr. and P. R. Taylor, J. Chem. Phys. **85**, 6510 (1986).

[35] C. W. Bauschlicher Jr. and P. R. Taylor, J. Chem. Phys. **86**, 1420 (1987).

[36] B. O. Roos and K. Andersson, Chem. Phys. Lett **245**, 215 (1995).

[37] K. L. Schuchardt et al., J. Chem. Inf. Model. **47**, 1045 (2007).

[38] T. H. Dunning Jr., J. Chem. Phys. **90**, 1007 (1989).

[39] A. J. H. Wachters, J. Chem. Phys. **52**, 1033 (1970).

[40] C. W. Bauschlicher Jr., S. R. Langhoff, and L. A. Barnes, J. Chem. Phys. **91**, 2399 (1989).

[41] C. Angeli, B. Bories, A. Cavallini, and R. Cimiraglia, J. Chem. Phys. **124**, 054108 (2006).

[42] P. Celani, H. Stoll, H.-J. Werner, and P. Knowles, Mol. Phys. **102**, 2369 (2004).

[43] B. O. Roos et al., J. Mol. Struct. THEOCHEM **388**, 257 (1996).

[44] K. Nakayama, H. Nakano, and K. Hirao, Int. J. Quantum Chem. **66**, 157 (1998).

[45] A. D. Becke, J.Chem.Phys. **98**, 5648 (1993).

[46] C. Lee, W. Yang, and R. Parr, Phys. Rev. B **37**, 785 (1988).

[47] R. McDiarmid, Chem. Phys. Lett. **188**, 423 (1992).

[48] J. P. Doering and R. McDiarmid, J. Chem. Phys. **73**, 3617 (1980).

[49] J. P. Doering and R. McDiarmid, J. Chem. Phys. **75**, 2477 (1981).

[50] O. A. Mosher, W. M. Flicker, and A. Kuppermann, Chem. Phys. Lett. **19**, 332 (1973).

[51] R. M. Gavin, S. Risemberg, and S. A. Rice, J. Chem. Phys. **58**, 3160 (1973).

[52] R. M. Gavin and S. A. Rice, J. Chem. Phys. **60**, 3231 (1974).

[53] A. Kuppermann, W. M. Flicker, and O. A. Mosher, Chem. Rev. **79**, 77 (1979).

[54] W. M. Flicker, O. A. Mosher, and A. Kuppermann, Chem. Phys. Lett. **45**, 492 (1977).

[55] M. Granville, G. Holtom, B. Kohler, R. Christensen, and K. D'Amico, J. Chem. Phys. **70**, 593 (1979).

[56] L. A. Heimbrook, B. E. Kohler, and I. J. Levy, J. Chem. Phys. **81**, 1592 (1984).

[57] D. G. Leopold, R. D. Pendley, J. L. Roebber, R. J. Hemley, and V. J. Vaida, J. Chem. Phys. **81**, 4218 (1984).

[58] L. D'Amico, C. Manos, and R. L. Christensen, J. Am. Chem. Soc. **102**, 1777 (1980).

CHAPTER 6

CANONICAL TRANSFORMATION THEORY : APPLICATION TO THE $2A_G^-$ AND $1B_U^+$ STATES IN POLYENES

(Some parts of this chapter are taken from Ref. [1].)

6.1 Introduction

Canonical transformation (CT) theory [2–4] is a recently developed method that is capable of handling the dynamic correlation in strongly correlated systems when applied on top of a multireference wavefunction, such as one obtained from complete active space self consistent field (CASSCF) or density matrix renormalization group (DMRG) calculations.

The accuracy and computational scaling of the theory has already been demonstrated on potential energy curves of small molecules such as H_2O, N_2, FeO and others [2–5]. When compared to other multireference dynamic correlation methods, it has been seen that the accuracy is decidedly superior to complete active space perturbation theory (CASPT2) and comparable to size-consistency corrected multireference configuration interaction (MRCI) methods. However, as the computational scaling ($O(n^6)$ where n is the number of electrons) is much better than that of the MRCI methods, CT theory is in principle tractable for much larger systems.

In this chapter, we applied the CT theory to the problem of low-lying electronic excitations in all-*trans* polyenes. The continuing relevance of polyene chemistry is due to the fact that polyene chromophores play a major role in biologically important photo-processes, ranging from light-harvesting to vision. All-*trans*

polyenes have C_{2h} point group symmetry. Additionally, polyenes have an approximate particle-hole symmetry (that is exact in the Hückel model) and states are often labeled by an approximate +/- particle-hole label. We shall henceforth restrict the discussion to the singlet states only. The ground state is $1A_g^-$, and the π excitations are variously of A_g^-, B_u^+, and B_u^- symmetry. The B_u^+ states are those with strong dipole-allowed transitions from the ground-state. The dark states are the A_g^- states (completely forbidden transitions from the ground state) and B_u^- states (weakly allowed transitions from the ground state). The essential problem is the position and number of dark states (below the $1B_u^+$ state) as a function of polyene length, which remains a matter of considerable debate [6–11].

The primary theoretical challenges are [12–21]:

1. From a valence-bond perspective the various excited states are very different in their "covalent" and "ionic" characters with varying amounts of dynamic and static correlation. The A_g^- and B_u^- states are predominantly composed of double excitations relative to the ground state while the B_u^+ states have more single excitation character. Reaching a balanced treatment is exceedingly difficult, and theoretical methods often give contradictory results.
2. Due to valence-Rydberg mixing (between the π^* and carbon $3p$), large basis sets are needed to obtain meaningful results.
3. When comparing to experiments, non-vertical excitation processes have to be considered. These require definite knowledge about the excited state geometry which is not trivial due to the presence of avoided crossings and a complicated potential surface.

Because of these many challenges, the correct description of polyene excited states from a theoretical perspective needs a multiconfigurational structure with dynamic correlation, along with large basis sets to include the possible mixing between valence and Rydberg states.

Short polyenes were first studied using semi-empirical Hamiltonians with a variety of methods such as full configuration interaction (FCI), MRCI, valence bond theory and effective valence shell Hamiltonians [22–24]. While much valuable insight has been provided by these calculations, the typical semi-empirical parametrization is biased towards the HOMO-LUMO ($1B_u^+$) transition. Using *ab-initio* Hamiltonians (with realistic basis sets) recent studies have employed coupled cluster (CC), multireference perturbation theories (MRMP and CASPT2) and multireference CI [13–21, 25–28]. CC methods tend to place predominantly "covalent" (i.e. dark) excited states too high in energy while MRMP and CASPT2 place such states too low. In our view, *ab-initio* MRCI calculations are probably the most reliable but due to their prohibitive cost have not been applied to molecules larger than butadiene. Our previous studies have shown that the typical accuracy of CT is comparable to that of MRCI, but the superior scalability allows us to address excited states even in the longer polyenes. This, together with the possibility of benchmarking the CT theory for excited states, is the basic motivation of the current work.

6.2 Theory

In this chapter, we will briefly review the CT theory. The details of the CT theory are given in Refs. [2, 3, 5].

CT theory is based on an exponential ansatz,

$$\Psi = e^{A}\Psi_0 \tag{6.1}$$

where Ψ_0 is a multiconfigurational wavefunction, generally obtained from a CASSCF calculation. A is an anti-Hermitian excitation operator $A^\dagger = -A$ that connects the active and external spaces. At the singles doubles (SD) level of the theory, A contains the single, double and semi-internal single excitations,

$$A = A_i^a(a_i^a - a_a^i) + A_{ij}^{ab}(a_{ij}^{ab} - a_{ab}^{ij}) + A_{ij}^{ak}(a_{ij}^{ak} - a_{ak}^{ij}). \tag{6.2}$$

The indices *ijk...* and *abc...* denote active and external indices respectively, and the excitation operators, $a_i^a = a_a^\dagger a_i$ and $a_{ij}^{ab} = a_a^\dagger a_b^\dagger a_j a_i$. Using the above notation, the energy and the amplitudes are obtained from

$$\langle\Psi_0|\bar{H}|\Psi_0\rangle = E \tag{6.3}$$

$$\langle\Psi_0|[\bar{H}, A]_{1,2}|\Psi_0\rangle = 0 \tag{6.4}$$

where $\bar{H}$ is the CT effective Hamiltonian defined by the Baker-Campbell-Hausdorff expansion,

$$\bar{H} = H + [H, A]_{1,2} + \frac{1}{2!}[[H, A]_{1,2}, A]_{1,2} + ... \tag{6.5}$$

The subscript $1, 2$ denotes an operator and cumulant decomposition (see Refs. [2, 3]) where the many body terms generated in the commutators are effectively reduced to 1 and 2 body operators. Because of the operator decomposition, $\bar{H}$ is a two body Hamiltonian just like H. The energy and amplitude expressions (6.4) require only 1 and 2 body active density matrices. (The fact that energy and amplitude expressions (6.4) require only 1 and 2 body active density matrices, make CT a two body density matrix functional theory and this is the basis of the connection to some recent work by Mazziotti on contracted Schrödinger equations [29].) Note that here we are applying the decompositions already at the

first (linear commutator). This defines the linear CT approximation (referred to in Ref. [5]), and all calculations in this chapter use this approximation, although we will not be using the "L" label for simplicity. The use of the operator and cumulant decomposition to simplify Eqns. 6.3 and 6.4 results in the favorable cost scaling of the CT theory ($n_{act}^2 n_{ext}^4$) [2, 3].

6.2.1 State averaged CT for excited states

In order to apply CT theory to excited states of the same symmetry in a stable numerical way, it is convenient to use a form of state averaging. Since the amplitude equations depend only on the 1 and 2 body density matrices γ_1, γ_2 (due to the operator and cumulant decomposition) we can formally write the amplitude equation (6.4) as a functional of these density matrices

$$R[\gamma_2] = 0 \Rightarrow A \tag{6.6}$$

(where we have omitted γ_1 since it can be obtained from γ_2). The corresponding state averaged amplitudes A^{sa} are defined from an amplitude equation where γ_2 is substituted by the state averaged γ_2^{sa} of the reference problem,

$$\gamma_{ij}^{kl(sa)} = \sum_m \langle \Psi_0^m | a_{ij}^{kl} | \Psi_0^m \rangle \tag{6.7}$$

where Ψ_0^m is the mth state of the reference system in the state averaged calculation. This gives

$$R[\gamma_2^{sa}] = 0 \Rightarrow A^{sa}. \tag{6.8}$$

From the state averaged A^{sa} we obtain the energies of the individual excited states by defining a *state specific* effective Hamiltonian $\bar{H}^m$ for each excited state using the amplitudes A^{sa} but the state specific density matrices γ_1^m, γ_2^m (defined

from the mth reference state) in the operator and cumulant decomposition

$$\bar{H}^m = H + [H, A^{sa}]_{1,2(m)} + \frac{1}{2!}[[H, A^{sa}]_{1,2(m)}, A^{sa}]_{1,2(m)} + ... \tag{6.9}$$

where the subscript $1, 2(m)$ denotes that the operator and cumulant decomposition is carried out with the density matrices γ_1^m, γ_2^m. When we take the expectation value of the state specific effective Hamiltonian $\bar{H}^m$ with the mth reference state, we get an appropriate CT energy for the mth excited state,

$$E^m = \langle \Psi_0^m | \bar{H}^m | \Psi_0^m \rangle. \tag{6.10}$$

6.2.2 Converging CT equations

There are a few numerical issues with the canonical transformation theory associated with converging the amplitude equations (6.4), in particular because the Jacobian associated with the amplitude equation can be nearly singular when using a multireference wavefunction. These convergence issues are discussed in detail in Refs. [3, 5, 30]. Here we use two of the methods for convergence described in [30].

Overlap truncation threshold method

In this method, spaces of excitation operators are truncated based on the overlap matrix of the first-order interacting space. Two truncation thresholds (with units of overlap) are used: τ_1 associated with the single excitations and semi-internal single excitations, and τ_2 associated with the double external excitations. Naturally, the thresholds should be chosen with care. There are two possible strategies. One is to choose the minimum possible threshold for which the

CT amplitude equations converge [5, 30]. However at this threshold, the energy is relatively sensitive to the changes in the threshold and further the convergence of the amplitude equations is very slow. Another strategy is to choose an appropriate threshold by probing all thresholds within reasonable limits (0.01 to 0.5) and finding a plateau in which the energy change is minimal. In this work, we have used the latter technique to define the appropriate thresholds.

Strongly contracted CT

The threshold technique of removing singularities in the Jacobian can be tedious since in principle we should carry out a number of CT calculations to decide on an appropriate threshold. In addition, it also requires diagonalization of the first-order interacting space overlap matrix, which in principle becomes prohibitively expensive ($O(n_{act}^9)$) with active spaces with more than 30 orbitals which arise in DMRG calculations. One way to bypass these problems is to remove singularities based on a physically motivated subset of the excitation degrees of freedom. This yields the strongly contracted variant of CT, which uses a small set of "strongly contracted" excitations or perturber functions, as introduced in n-electron valence perturbation theory (NEVPT2) [31, 32], to span the first-order interacting space. The overlap matrix of the perturber functions is then diagonal by construction. While even within this smaller parametrization, there can still be some small eigenvalues of the Jacobian, generally these are much more easily removed than for the full Jacobian. A single universal threshold ($0.1E_h$) is used in all calculations, corresponding to a cutoff of the smallest perturber function excitation energy as measured with the Dyall Hamiltonian [30].

6.3 Computational methodology

6.3.1 Overview

Starting from a complete active space self-consistent field (CASSCF) calculation in a π-active space (of varying size) and with large basis sets saturated with Rydberg functions, we have performed canonical transformation calculations using the ORZ code on the all-*trans* polyenes, C_4H_6, C_6H_8 and C_8H_{10} at their equilibrium geometries, using both the overlap threshold method (CTSD thresh) and strong contraction method (sc-CTSD) to converge the amplitude equations. For comparison, CASPT2 (the partially uncontracted `rs2c` variant implemented in MOLPRO) [33, 34], MRCI (with and without the Q Davidson correction) [35, 36] and multireference averaged coupled pair functional (MRACPF) [37–39] calculations were carried out using the MOLPRO package [40]. We now give further details on the basis set, active space, and geometries used.

6.3.2 Basis set

To account for the possibility of valence-Rydberg mixing in the $1B_u^+$ states, we need to include additional diffuse p functions in the carbon atomic basis. Following Ref. [25] we started from Dunning's cc-pVDZ [41] basis and augmented the carbon basis with up to 3 additional p functions, to saturate the p basis. The p exponents used were 0.04041, 0.01080 and 0.00250 (the first two are the p exponents found in the d-aug-cc-pVDZ basis). From now on we will refer to the carbon basis sets thus formed (with one, two or three additional functions) as aug′-cc-pVDZ ($3s3p1d$), d-aug′-cc-pVDZ ($3s4p1d$) and t-aug′-cc-pVDZ ($3s5p1d$)

where the ′ indicates that the basis sets are augmented only in their p functions. Dunning's cc-pVDZ basis was used for hydrogen.

6.3.3 Active space

To take into account the multireference nature of the valence-Rydberg mixing, we used both $\pi(n, n)$ and "double" $\pi(n, 2n)$ active spaces where n is the number of C atoms. The double π space allows Rydberg $3p$ configurations to mix with the valence states in a multireference way.

6.3.4 Geometry

The polyenes have C_{2h} symmetry and were oriented in the xy plane, i.e. with the C_2 axis as the z axis. The molecular geometries were optimized at the density functional level using the B3LYP functional [42, 43] as implemented in GAUSSIAN03 [44]. The basis set used for the optimization was Dunning's cc-pVDZ basis [41].

6.4 Results and discussions

6.4.1 Butadiene

The low-lying electronic spectrum of butadiene has been intensively studied both experimentally and theoretically. Owing to many Rydberg peaks and pos-

sible excited state geometrical relaxation, the spectrum is very broad and there have been many different interpretations [6, 15, 19–21]. Theoretical predictions also greatly vary depending on the level of correlation and the basis sets used [13, 15, 23, 45–52].

Effect of active space and basis set

To gain some understanding of the problem, the effects of active space and basis set were first probed at the CASSCF and MRCI level of correlation with the cc-pVDZ and augmented basis sets aug′-cc-pVDZ, d-aug′-cc-pVDZ and t-aug′-cc-pVDZ described earlier. Table 6.1 gives the $1B_u^+$ excitation energies using the (4e, 4o) and (4e, 8o) active spaces with various basis sets. While there is some change in the $1B_u^+$ excitation energy when doubling the active space size using the valence basis cc-pVDZ, there is a significant decrease in the excitation energies when using the augmented cc-pVDZ basis sets which contain Rydberg functions. (In the CASSCF level the decrease when doubling the active space is 0.42 eV in the cc-pVDZ basis and 0.69 eV in the most augmented basis, while the MRCI decrease in excitation energies goes from 0.14 eV to 0.35 eV.) This follows from our general understanding that the considerable valence-Rydberg mixing in butadiene needs both valence and Rydberg states to be present in the active space. We also observe that the effect of increased active space is much greater at the CASSCF level of theory than at the MRCI level and we would expect the effect to be smaller still with corrected MRCI methods, such as MRACPF, which incorporate even more correlation out of the active space.

The excitation energy converges from above when augmenting the basis but appears to reach a plateau at the doubly augmented level. This indicates that

Table 6.1: G.S. $\rightarrow 1B_u^+$ excitation energy (in eV) with CASSCF and MRCI methods for butadiene with different active spaces and basis sets. The expectation value $\langle x^2 \rangle$ (in $Bohr^2$) is a measure of the diffuseness and therefore the Rydberg nature of the excited $1B_u^+$ state.

Method and Active space	DZp [1]	$\langle x^2 \rangle$	DZp+p [2]	$\langle x^2 \rangle$	DZp+2p [3]	$\langle x^2 \rangle$	DZp+3p [4]	$\langle x^2 \rangle$
CASSCF(4,4)	8.12	22.21	7.56	31.87	7.08	73.74	7.08	74.11
CASSCF(4,8)	7.70	22.25	6.92	33.42	6.39	69.54	6.39	69.33
MRCI(4,4)	7.17	22.09	7.01	31.05	7.02	72.26	7.02	72.62
MRCI(4,8)	7.03	22.12	6.77	32.35	6.67	67.83	6.67	67.62

[a] DZp = cc-pVDZ
[b] DZp+p = aug′-cc-pVDZ
[c] DZp+2p = d-aug′-cc-pVDZ
[d] DZp+3p = t-aug′-cc-pVDZ

the d-aug′-cc-pVDZ basis is probably the smallest carbon basis set that should be used for the butadiene spectrum. One more point can be noted from the table. The expectation value $\langle x^2 \rangle$ which defines the extent of the wavefunction or its diffuseness also increases with augmenting the basis set. With these basis sets and active spaces, the $1B_u^+$ state is probably too diffuse, especially at the CASSCF level. Previous studies have found that the diffuseness can be reduced by appropriate state averaging, but we have not performed such calculations here [26].

Butadiene excitation energies from canonical transformation theory

In table 6.2, CASPT2, MRCI, MRCI with Davidson correction (MRCI+Q) and MRACPF excitation energies for the $1B_u^+$ state are presented alongside CT excitation energies for this state using both the overlap truncation threshold (CTSD thresh) and strongly contracted (sc-CT) variants.

Especially for the larger basis sets, there is reasonable agreement between the different methods: at the t-aug′-cc-pVDZ level the theoretical excitation energies range from 6.60 eV (CASPT2(4,4)) to 7.02 eV (MRCI(4,4)). As expected the CASPT2 excitation energies are considerably lower than those obtained from MRCI and its variants. As we regard the MRCI calculations to be more reliable we find it encouraging that the CT excitation energies are quite comparable to those obtained from MRCI type methods. (Compared to the literature, we should note that there is some variation in the reported CASPT2 and MRCI results. This is because the $1B_u^+$ state is quite sensitive both to the composition of the active space and the kind of basis used. Reported CASPT2/MRMP $1B_u^+$ excitation energies range from 6.48 eV [27] to 6.12 eV [49], while those for MRCI

Table 6.2: $1B_u^+$ excitation energies (in eV) for butadiene with various correlation methods and basis sets. The active spaces used are (4e, 4o) and (4e, 8o). The τ_1 and τ_2 reported correspond to the CT truncation thresholds used for the particular calculations.

Method and Active space	DZp[1]	DZp+p[2]	DZp+2p[3]	DZp+3p[4]
CASSCF(4,4)	8.12	7.56	7.08	7.08
CASSCF(4,8)	7.70	6.92	6.39	6.39
CASPT2$(4,4)$	6.32	6.14	6.60	6.60
CASPT2$(4,8)$	6.47	6.45	6.65	6.71
MRCI$(4,4)$	7.17	7.01	7.02	7.02
MRCI$(4,8)$	7.03	6.77	6.67	6.67
MRCI+Q$(4,4)$	6.73	6.67	6.88	6.88
MRCI+Q$(4,8)$	6.70	6.70	6.71	6.71
ACPF$(4,4)$	6.71	6.46	6.62	6.62
ACPF$(4,8)$	6.67	6.62	6.72	6.72
CTSD thresh$(4,4)$	6.93[6]	$n.c.$[5]	$n.c.$	$--$
CTSD thresh$(4,8)$	7.02[6]	6.82[7]	6.81[8]	$--$
sc-CTSD$(4,4)$	6.86	$n.c.$	$n.c.$	$--$
sc-CTSD$(4,8)$	7.05	6.75	6.72	$--$
Expt.	5.92[9]			

[1]DZp = cc-pVDZ. [2]DZp+p = aug′-cc-pVDZ.
[3]DZp+2p = d-aug′-cc-pVDZ. [4]DZp+3p = t-aug′-cc-pVDZ.
[5]n.c. = not converged.
[6]$\tau_1 = 0.15$, $\tau_2 = 0.05$.
[7]$\tau_1 = 0.20$, $\tau_2 = 0.05$.
[8]$\tau_1 = 0.25$, $\tau_2 = 0.05$.
[9][8, 53, 54].

Table 6.3: $2A_g^-$ excitation energies (in eV) for butadiene with various correlation methods and basis sets. The active spaces used are (4e, 4o) and (4e, 8o).The τ_1 and τ_2 reported correspond to the CT truncation thresholds used for the particular calculations.

Method and Active space	DZp[1]	DZp+p[2]	DZp+2p[3]	DZp+3p[4]
CASSCF(4,4)	6.65	6.59	6.59	6.59
CASSCF(4,8)	6.69	6.54	6.54	6.54
CASPT2$(4,4)$	6.48	6.38[5]	6.45[6]	6.45[6]
CASPT2$(4,8)$	6.47	6.28	6.35[5]	6.38[6]
MRCI$(4,4)$	6.70	6.67	6.67	6.67
MRCI$(4,8)$	6.66	6.60	6.53	6.53
MRCI+Q$(4,4)$	6.67	6.63	6.63	6.63
MRCI+Q$(4,8)$	6.63	6.58	6.50	6.50
CTSD thresh$(4,4)$	6.51[8]	$n.c.$[7]	$n.c.$	$--$
CTSD thresh$(4,8)$	6.66[9]	6.49[10]	$n.c.$	$--$
sc-CTSD$(4,4)$	6.55	$n.c.$	$n.c.$	$--$
sc-CTSD$(4,8)$	6.72	6.59	6.58	$--$
Expt.	$5.4-7.3$[11]			

[1]DZp = cc-pVDZ.
[2]DZp+p = aug′-cc-pVDZ.
[3]DZp+2p = d-aug′-cc-pVDZ.
[4]DZp+3p = t-aug′-cc-pVDZ.
[5]level shift = 0.2 $a.u.$
[6]level shift = 0.25 $a.u.$
[7]n.c. = not converged.
[8] τ_1 = 0.2, τ_2 = 0.05.
[9] τ_1 = 0.2, τ_2 = 0.05.
[10] τ_1 = 0.3, τ_2 = 0.05.
[11][8, 53, 55, 56].

range from 7.05 eV [14] to 6.23 eV [45]. Our $1B_u^+$ excitation energies are on the high side, since our active space seems to favor a Rydberg-like state.) Despite the general agreement between the methods, there is quite different behavior with respect to the basis set and active space dependence. In particular, CASPT2 appears to have a fairly strong basis set dependence and the excitation energy converges from below, while MRCI type methods have only a weak basis set dependence. This reflects the different balance of valence and Rydberg states at these two levels of approximation. While CT excitation energies converge from above much like the MRCI-type methods, convergence problems with the CT theories prevent application to the largest t-aug′-cc-pVDZ basis. The sc-CT calculations converge much more easily than the overlap truncation CT calculations, at least at the truncation thresholds considered here.

Compared to the estimated experimental "vertical" excitation energy of 5.92 eV, all the theoretical predictions seem too high. This, however could be due to the excited state geometry relaxation which has been proposed as one of the reason for the broad experimental excitation spectrum. Previous theoretical studies [25, 47, 51, 57] show that allowing the geometry of the $1B_u^+$ excited state to relax and become out of plane leads to a much better agreement between the theoretical and experimental $1B_u^+$ excitation energies.

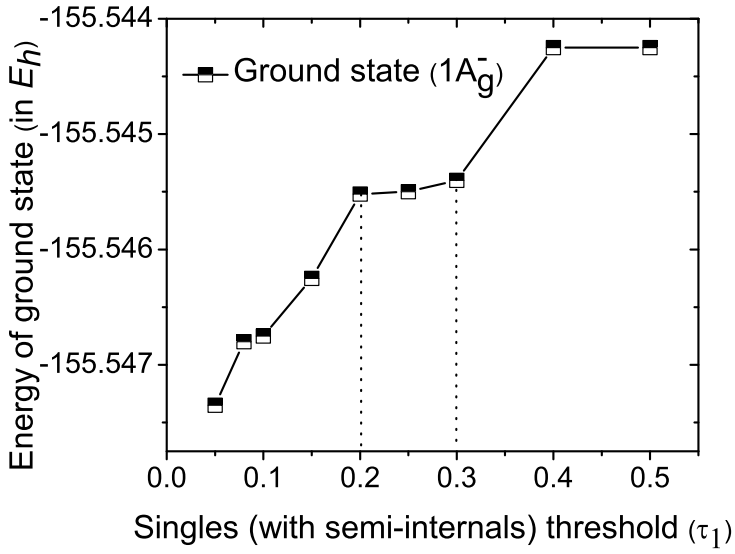
While the CT excitation energies are broadly comparable to the MRCI results, the CT theory allows application to much larger active spaces. We have taken advantage of this to examine the effect of incorporating the σ framework into the active space. By performing a sc-CT calculation with a (10e, 14o) active space that includes the 3 bonding and antibonding carbon framework σ orbitals, we find that the $1B_u^+$ excitation energy changes by only 0.10 eV. This suggests that σ

framework effects can be adequately captured at the dynamic correlation level.

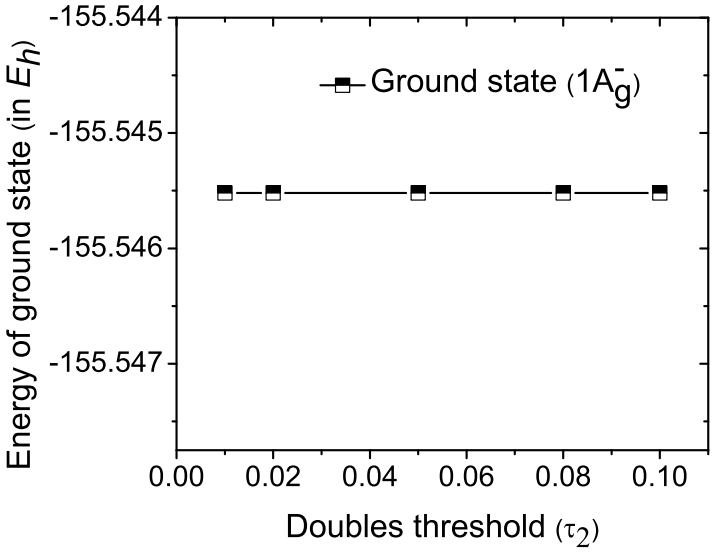
Table 6.3 shows the analogous calculations for the $2A_g^-$ state. These results were obtained by state averaging over the $1A_g^-$ and $2A_g^-$ states. (This could not be carried out with the MRACPF method implemented in MOLPRO and so those results are missing from the table). The effect of active space and basis sets is much less for the $2A_g^-$ state than for the $1B_u^+$ state since dynamic correlation is less important in this state and there is no single-particle valence-Rydberg mixing (for symmetry reasons). All methods predict that the $2A_g^-$ state is very close in energy to the $1B_u^+$ state, 0.47 eV being the largest separation at the MRCI(4,4) level. Once again, the CT calculations yield results quite comparable to those from MRCI methods. Because of the different basis set dependence of the energy of the two states, the ordering of the states can change with basis set size. With the largest basis sets, however, all methods seem to predict that the $2A_g^-$ lies below the $1B_u^+$ state, although as mentioned above, our $1B_u^+$ excitation energies are probably a bit high. Overall our $2A_g^-$ excitation energies are in good agreement with recent MRCI predictions [13, 26].

CT truncation thresholds

Since the overlap truncation threshold method for converging the CT equations yields results that depend in principle on the threshold, here we show a threshold analysis for the butadiene molecule. Figs. 6.1 and 6.2 illustrate the threshold dependence of the CT total energy for the ground state and $1B_u^+$ states (using the aug'-cc-pVDZ basis for C) while Fig. 6.3 shows the dependence of the $1B_u^+$ excitation energy on threshold. For the singles and semi-internal singles truncation threshold τ_1 we see that there is a plateau around 0.2-0.3 in τ_1, and thus for

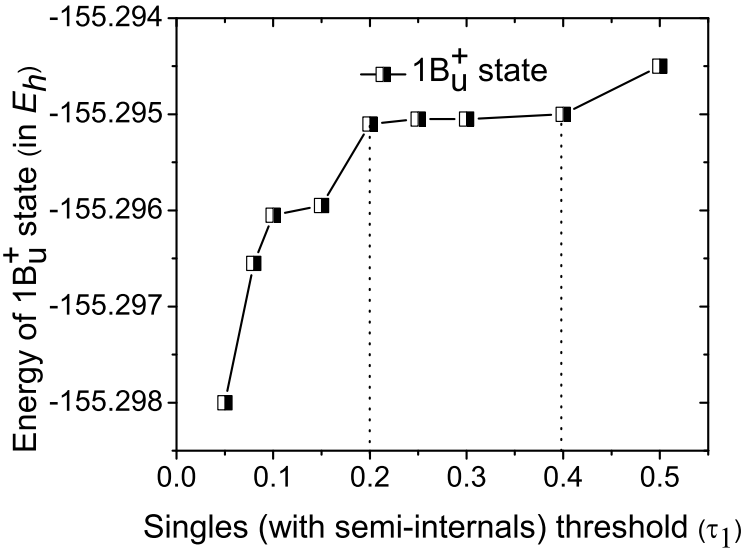


(a) τ_1 (keeping $\tau_2 = 0.05$).

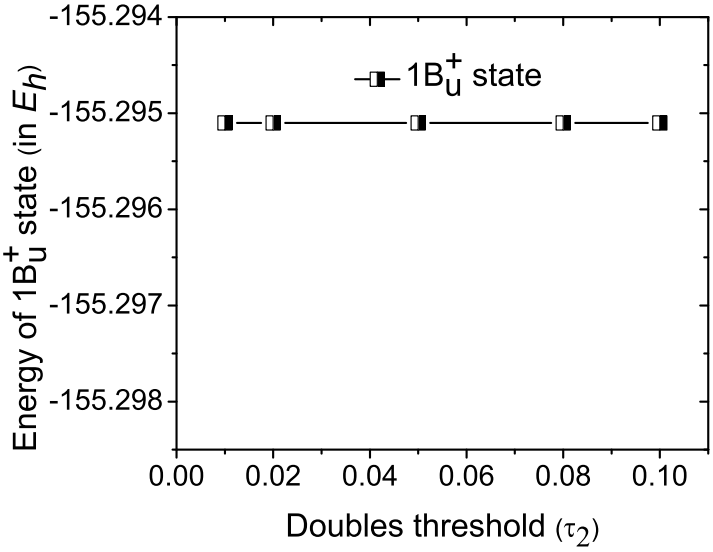


(b) τ_2 (keeping $\tau_1 = 0.25$).

Figure 6.1: Threshold analysis of the ground state total energy of C_4H_6 with aug′-cc-pVDZ basis set. The energies are in E_h. The dotted lines in the plot of energy vs. τ_1 mark the plateau region where the change in energy with threshold change is minimal. The plot of energy vs. τ_2 shows that the energy is insensitive to changes in τ_2 in the range $0.01 - 0.1$.



(a) τ_1 (keeping $\tau_2 = 0.05$).



(b) τ_2 (keeping $\tau_1 = 0.25$).

Figure 6.2: Threshold analysis of the total energy of $1B_u^+$ state of C_4H_6 with aug′-cc-pVDZ basis set. The energies are in E_h. The dotted lines in the plot of energy vs. τ_1 mark the plateau region where the change in energy with threshold change is minimal. The plot of energy vs. τ_2 shows that the energy is insensitive to changes in τ_2 in the range $0.01 - 0.1$.

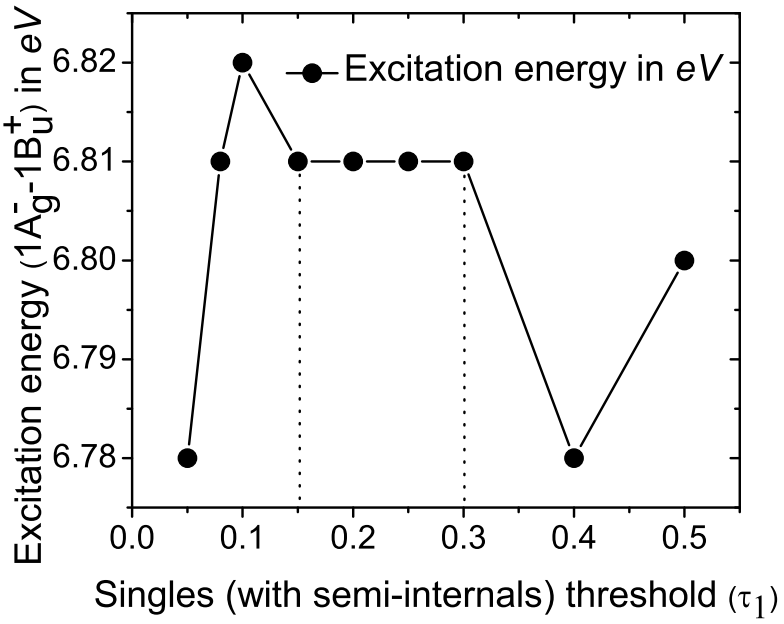


Figure 6.3: Dependence of the $1A_g^- \rightarrow 1B_u^+$ excitation energy of C_4H_6 on the CT truncation threshold used. Since both ground state and $1B_u^+$ excited state exhibit plateaus in the region $\tau_1 = 0.2 - 0.3$, a similar plateau is observed in the excitation energy. The value of τ_2 is kept fixed at 0.05.

this state, we used $\tau_1 = 0.2$. The energies are much less sensitive to the doubles threshold τ_2: this could be kept constant throughout all the calculations with $\tau_2 = 0.05$. For the other active spaces, states and basis sets, a similar threshold analysis was carried out to determine reasonable values of τ_1 which were reported in tables 6.2 and 6.3.

While the results of the threshold truncated CT (see previous section) appear quite reasonable it is clear that there can be situations when the determination of an appropriate threshold is not simple. Furthermore threshold analysis involves a number of tedious calculations. For this reason in the higher polyenes we used only the strongly contracted variant of CT to circumvent the kind of threshold

analysis described here.

6.4.2 Longer polyenes

Table 6.4: $1B_u^+$ excitation energies (in eV) for hexatriene with various correlation methods and basis sets. The active spaces used are (6e, 6o) and (6e, 12o). MRCI and MRCI+Q methods are carried out with only (6e, 6o) active space.

Method and Active space	DZp[1]	DZp+p[2]	DZp+2p[3]
CASSCF$(6,6)$	7.40	7.28	7.28
CASSCF$(6,12)$	6.71	6.37	6.21
CASPT2$(6,6)$	5.07	4.83	4.83
CASPT2$(6,12)$	5.46	5.70	6.22[4]
MRCI$(6,6)$	6.42	6.28	6.27
MRCI+Q$(6,6)$	5.95	5.80	5.79
sc-CTSD$(6,6)$	5.90	5.75	5.78
sc-CTSD$(6,12)$	6.03	6.08	6.09
Expt.	$4.93-5.2$[5]		

[1]DZp = cc-pVDZ.
[2]DZp+p = aug′-cc-pVDZ.
[3]DZp+2p = d-aug′-cc-pVDZ.
[4]level shift = 0.1 $a.u.$
[5][58, 59].

Analogously to the preceding calculations on butadiene, we computed vertical excitation energies to the lowest $2A_g^-$ and $1B_u^+$ states using CT (with only the strongly contracted variant) and a variety of multireference quantum chemical methods in all-*trans* hexatriene and all-*trans* octatetraene. In these larger

Table 6.5: $2A_g^-$ excitation energies (in eV) for hexatriene with various correlation methods and basis sets. The active spaces used are (6e, 6o) and (6e, 12o). MRCI and MRCI+Q methods are carried out with only (6e, 6o) active space.

Method and Active space	DZp[1]	DZp+p[2]	DZp+2p[3]
CASSCF$(6,6)$	5.55	5.48	5.48
CASSCF$(6,12)$	5.55	5.52	5.53
CASPT2$(6,6)$	5.21	5.21	5.28
CASPT2$(6,12)$	5.23	5.28	5.32
MRCI$(6,6)$	5.57	5.55	5.55
MRCI+Q$(6,6)$	5.51	5.51	5.53
sc-CTSD$(6,6)$	5.57	5.54	5.55
sc-CTSD$(6,12)$	5.55	5.50	5.50
Expt.	$5.7-6.45$[4]		

[1]DZp = cc-pVDZ.
[2]DZp+p = aug′-cc-pVDZ.
[3]DZp+2p = d-aug′-cc-pVDZ.
[4][59].

systems some new considerations come into play in the quantum chemistry. For example, already for the hexatriene molecule, the steep cost scaling of the MRCI methods prevents the use of a double π-active space, and in octatetraene, the MRCI calculations simply were not feasible in a reasonable amount of time. This limitation of course does not apply to the CT theory, with its n^6 (where n is system size) computational scaling.

Tables 6.4 and 6.5 show the excitation energies for the various methods for the hexatriene molecule. Examining first the $1B_u^+$ state of hexatriene, there is ar-

Table 6.6: $1B_u^+$ excitation energies (in eV) for octatetraene with various correlation methods and basis sets. The active spaces used are (8e, 8o) and (8e, 16o).

Method and Active space	DZp[1]	DZp+p[2]
CASSCF$(8,8)$	6.66	6.56
CASSCF$(8,16)$	6.16	5.85
CASPT2$(8,8)$	4.40	4.41[3]
CASPT2$(8,16)$	4.72	5.29[3]
sc-CTSD$(8,8)$	5.26	$n.c.$[4]
sc-CTSD$(8,16)$	5.25	5.29
Expt.	4.4[5]	

[1]DZp = cc-pVDZ.
[2]DZp+p = aug′-cc-pVDZ.
[3]level shift = 0.2 $a.u.$
[4]n.c. = not converged.
[5][60].

guably a weaker basis set dependence than in butadiene: all methods seem to saturate at the aug′-cc-pVDZ level except for CASPT2(6,12) which appears to have an intruder state problem for the largest basis set. This would be consistent with a smaller amount of valence-Rydberg mixing due to the decreased HOMO-LUMO gap. Unlike in butadiene, the MRCI+Q excitation energies appear quite different from those of MRCI. We might speculate that this is related to the lack of size-extensivity of MRCI in this larger molecule. Certainly the sc-CT (which is size-extensive) excitation energies match very closely the MRCI+Q energies (in the same active space). There is also much greater disagreement compared to butadiene between the CASPT2 excitation energies and the MRCI/CT excita-

Table 6.7: $2A_g^-$ excitation energies (in eV) for octatetraene with various correlation methods and basis sets. The active spaces used are (8e, 8o) and (8e, 16o).

Method and Active space	DZp[1]	DZp+p[2]
CASSCF$(8,8)$	4.72	4.71
CASSCF$(8,16)$	4.75	4.71
CASPT2$(8,8)$	4.35	4.34[3]
CASPT2$(8,16)$	4.37	4.47[3]
sc-CTSD$(8,8)$	4.60	$n.c.$[4]
sc-CTSD$(8,16)$	4.72	4.72
Expt.	3.54[5]	

[1]DZp = cc-pVDZ.
[2]DZp+p = aug′-cc-pVDZ.
[3]level shift = 0.2 $a.u.$
[4]n.c. = not converged.
[5][61, 62].

tion energies for the $1B_u^+$ state with CASPT2 placing the $1B_u^+$ state much lower in energy. While the agreement with experiment appears best for CASPT2, we note that the experimental energies are almost certainly not vertical in this system. Given the observed intruder state problem for CASPT2 in this molecule as well as the much higher placement of the level by MRCI, MRCI+Q and CT theories, we consider it likely that the CASPT2 $1B_u^+$ state is simply too low. Looking at the $2A_g^-$ state in hexatriene, we see a repeat of the butadiene situation, where the effect of basis set, active space and correlation method is less significant than in the $1B_u^+$ state. The sc-CT results closely follow those of MRCI+Q, while the CASPT2 excitation energies are somewhat smaller. All lie in reasonable agree-

ment ($\approx 5.5eV$ as compared to the lowest experimental value of 5.7 eV) with the very wide range of reported experimental $2A_g^-$ excitation energies and all methods predict that the $2A_g^-$ vertical excitation lies below the $1B_u^+$ vertical excitation.

Tables 6.7 and 6.6 give the excitation energies for the octatetraene molecule. As discussed previously, only CASPT2 and sc-CT calculations were feasible for this system. Convergence in this system is non-trivial: all the CASPT2 calculations required a level shift of 0.2 $a.u.$, while the sc-CT calculation in the smaller $(8e, 8o)$ active space could not be converged for the augmented basis set. However, as far as we are aware the CT calculations represent the first time that dynamic correlation beyond second order perturbation theory has been used to study the spectrum of this molecule. We see results broadly consistent with what was observed in hexatriene, namely that the CASPT2 excitation energies are lower than those obtained using sc-CT. (The 5.29 eV excitation energy for the augmented basis with double π active space with CASPT2 may be due to an intruder state problem since we needed to use a level shift.) Both are considerably higher than the experimental values, but this is almost certainly due to excited state relaxation.

6.5 Conclusion

In this chapter, CT theory has been applied to study the lowest two singlet excited states of the all-*trans* polyenes, butadiene, hexatriene, and octatetraene. These calculations employed both large active spaces ("single" and "double" π active spaces) as well as realistic Rydberg augmented basis sets. In all the cases

examined, the CT theory excitation energies closely resemble those from size-consistency corrected multireference configuration interaction theory, but the superior cost scaling of CT theory allows it to be applied to considerably larger problems, in this case, the excited states of octatetraene.

While discrepancy remains between theory and experiment for the $2A_g^-$ and $1B_u^+$ states in all the polyenes considered, the convergence of the different theoretical methods for butadiene suggests that the primary source of the errors is likely to be non-vertical character in the experimental transitions. The study of non-vertical processes is naturally computationally intensive, and we believe that the accuracy and favorable scaling of CT theory will make it a useful tool for this problem, particularly for larger systems that are not currently treatable by other high-level methods.

BIBLIOGRAPHY

[1] D. Ghosh, E. Neuscamman, T. Yanai, and G. K.-L. Chan, J. Chem. Theor. Comput. (submitted).

[2] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **124**, 194106 (2006).

[3] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **127**, 104107 (2007).

[4] G. K.-L. Chan and T. Yanai, Adv. Chem. Phys. **134**, 343 (2007).

[5] E. Neuscamman, T. Yanai, and G. K.-L. Chan, J. Chem. Phys. **130**, 124102 (2009).

[6] R. McDiarmid and A.-H. Sheybani, J. Chem. Phys. **89**, 1255 (1988).

[7] D. G. Leopold, R. D. Pendley, J. L. Roebber, R. J. Hemley, and V. Vaida, J. Chem. Phys. **81**, 4218 (1984).

[8] J. P. Doering and R. McDiarmid, J. Chem. Phys. **73**, 3617 (1980).

[9] K. L. D'Amico, C. Manos, and R. L. Christensen, J. Am. Chem. Soc. **102**, 1777 (1980).

[10] F. Zerbetto and M. Z. Zgierski, J. Chem. Phys. **93**, 1235 (1990).

[11] A. Sabljić and R. McDiarmid, J. Chem. Phys. **84**, 2062 (1986).

[12] M. Aoyagi, Y. Osamura, and S. Iwata, J. Chem. Phys. **83**, 1140 (1990).

[13] L. Serrano-Andrés, J. Sánchez-Mańn, and I. Nebot-Gil, J. Chem. Phys. **97**, 7499 (1992).

[14] R. P. Hosteny, T. H. Dunning Jr., R. R. Gilman, A. Pipano, and I. Shavitt, J. Chem. Phys. **62**, 4764 (1975).

[15] R. J. Cave, J. Chem. Phys. **92**, 2450 (1990).

[16] R. J. Cave and E. R. Davidson, J. Phys. Chem. **92**, 614 (1988).

[17] B. R. Brooks and H. F. Schaefer III, J. Chem. Phys. **68**, 4839 (1978).

[18] C. Petrongolo, R. J. Buenker, and S. D. Peyerimhoff, J. Chem. Phys. **76**, 3655 (1982).

[19] J. Lappe and R. J. Cave, J. Phys. Chem. A **104**, 2294 (2000).

[20] A. C. Lasaga, R. J. Aerni, and M. Karplus, J. Chem. Phys. **73**, 5230 (1980).

[21] V. Bachler and K. Schaffner, Chem. Eur. J. **6**, 959 (2000).

[22] P. Tavan and K. Schulten, J. Chem. Phys. **70**, 5407 (1978).

[23] R. L. Graham and K. F. Freed, J. Chem. Phys. **96**, 1304 (1992).

[24] Z. G. Soos and S. Ramasesha, J. Chem. Phys. **90**, 1067 (1989).

[25] M. Boggio-Pasqua, M. J. Bearpark, M. Klene, and M. A. Robb, J. Chem. Phys. **120**, 7849 (2004).

[26] M. Dallos and H. Lischka, Theor. Chem. Acc. **112**, 16 (2004).

[27] K. Nakayama, H. Nakano, and K. Hirao, Int. J. Quantum Chem. **66**, 157 (1998).

[28] Y. Kurashige, H. Nakano, Y. Nakao, and K. Hirao, Chem. Phys. Lett. **400**, 425 (2004).

[29] A. E. DePrince III and D. A. Mazziotti, J. Chem. Phys. **127**, 104104 (2007).

[30] E. Neuscamman, T. Yanai, and G. K.-L. Chan, J. Chem. Phys. (submitted).

[31] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, J. Chem. Phys. **114**, 10252 (2001).

[32] C. Angeli, R. Cimiraglia, and J.-P. Malrieu, J. Chem. Phys. **117**, 9138 (2002).

[33] P. Celani and H.-J. Werner, J. Chem. Phys. **112**, 5546 (2000).

[34] H.-J. Werner, Mol. Phys. **89**, 645 (1996).

[35] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988).

[36] P. J. Knowles and H.-J. Werner, Chem. Phys. Lett. **145**, 514 (1988).

[37] H.-J. Werner and P. Knowles, Theor. Chim. Acta **78**, 175 (1990).

[38] R. J. Gdanitz and R. Ahlrichs, Chem. Phys. Lett. **143**, 413 (1988).

[39] P. G. Szalay and R. J. Bartlett, Chem. Phys. Lett. **214**, 481 (1993).

[40] H.-J. Werner et al., Molpro, version 2006.1, a package of ab initio programs, 2006, see http://www.molpro.net.

[41] T. H. Dunning Jr., J. Chem. Phys. **90**, 1007 (1989).

[42] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[43] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[44] M. J. Frisch et al., GAUSSIAN 03, REVISION C.02, Gaussian, Inc., Wallingford CT, 2004, see http://www.gaussian.com/.

[45] R. J. Cave and E. R. Davidson, J. Phys. Chem. **91**, 4481 (1987).

[46] S. Shih, R. J. Buenker, and S. D. Peyerimhoff, Chem. Phys. Lett. **16**, 244 (1976).

[47] P. G. Szalay, A. Karpfen, and H. Lischka, Chem. Phys. **130**, 219 (1989).

[48] M. A. C. Nascimento and W. A. Goddard III, Chem. Phys. **36**, 147 (1980).

[49] L. Serrano-Andrés, M. Merchán, I. Nebot-Gil, R. Lindh, and B. O. Roos, J. Chem. Phys. **98**, 3151 (1993).

[50] O. Kitao and H. Nakatsuji, Chem. Phys. Lett. **143**, 528 (1988).

[51] A. C. Lasaga, R. J. Aerni, and M. Karplus, J. Chem. Phys. **73**, 5230 (1980).

[52] P. G. Szalay and R. J. Bartlett, J. Chem. Phys. **101**, 4936 (1994).

[53] R. McDiarmid, J. Chem. Phys. **64**, 514 (1976).

[54] O. A. Mosher, W. M. Flicker, and A. Kuppermann, J. Chem. Phys. **59**, 6502 (1973).

[55] R. McDiarmid, Chem. Phys. Lett. **34**, 130 (1975).

[56] L. J. Rothberg, D. P. Gerrity, and V. Vaida, J. Chem. Phys. **73**, 5508 (1980).

[57] V. Bonaćić-Koutecký, M. Persico, D. Döhnert, and A. Sevin, J. Am. Chem. Soc. **104**, 6900 (1982).

[58] W. M. Flicker, O. A. Mosher, and A. Kuppermann, Chem. Phys. Lett. **45**, 492 (1977).

[59] R. M. Gavin Jr. and S. A. Rice, J. Chem. Phys. **60**, 3231 (1974).

[60] L. A. Heimbrook, B. E. Kohler, and I. J. Levy, J. Chem. Phys. **81**, 1592 (1984).

[61] M. F. Granville, G. R. Holtom, B. E. Kohler, R. L. Christensen, and K. L. D'Amico, J. Chem. Phys. **70**, 593 (1979).

[62] M. F. Granville, G. R. Holtom, and B. E. Kohler, J. Chem. Phys. **72**, 4671 (1980).

CHAPTER 7

APPLICATION OF DMRG-CASSCF AND DYNAMIC CORRELATION METHODS TO THE CAROTENOID EXCITED STATES

7.1 Introduction

Photosynthesis is the process of converting sunlight into chemical energy which occurs in photosynthetic bacteria, cyanobacteria, algae and green plants. During photosynthesis green plants use the sunlight as the source of energy and convert simple raw materials like carbon dioxide and water into sugar, carbohydrates, protein and lipids, all of which serve as food for living organisms.

Carotenoids are one of the major groups of accessory pigments involved in absorption of light in photosynthetic centers apart from the chlorophyll itself. They are long chained conjugated hydrocarbons distinguished from one another by their end groups. Carotenoids absorb light in the region of ultraviolet to about 550 nm in the visible range depending on their degree of conjugation as well as the type of end groups, in particular the presence or absence of oxygen or hydroxy groups.

The two major functions of carotenoids in the photosynthetic center are [1, 2]:

1. Light harvesting

   While light harvesting the carotenoid molecules absorb visible light (near the green region which is not effectively absorbed by the chlorophyll), and are excited into a singlet excited state. This excited singlet then transfers energy to the chlorophyll or bacteriochlorophyll. Therefore, in the light

harvesting process the light energy (in certain regions of the spectrum) is absorbed by the carotenoids, and rapidly transferred to the chlorophyll reaction center to perform photochemistry. The mechanism of this fast transfer depends on the energy of the excited states of the carotenoids as also their relative orientation and distance with the chlorophyll molecules. The light harvesting process in photosynthesis can be written as

$$\begin{aligned} \mathrm{Car} + h\nu &\rightarrow {}^1\mathrm{Car}^* \\ {}^1\mathrm{Car}^* + \mathrm{Chl} &\rightarrow \mathrm{Car} + {}^1\,\mathrm{Chl}^*, \end{aligned} \tag{7.1}$$

where Chl refers to the chlorophyll or bacteriochlorophyll molecules and Car refers to the carotenoid molecules.

2. Photo protection

The photo protective role of carotenoids is through the mediation of the triplet excited states of carotenoids,

$$\begin{aligned} \mathrm{Chl} + h\nu &\rightarrow {}^1\mathrm{Chl}^* \\ {}^1\mathrm{Chl}^* &\rightarrow {}^3\mathrm{Chl}^* \\ {}^3\mathrm{Chl}^* + \mathrm{Car} &\rightarrow {}^3\mathrm{Car}^* + \mathrm{Chl}. \end{aligned} \tag{7.2}$$

Carotenoids prevents the formation of the singlet oxygen or rapidly quench the singlet oxygen if formed, as given by the following equations,

$$\begin{aligned} {}^3\mathrm{Chl}^* + \mathrm{O}_2 &\rightarrow \mathrm{O}_2^*({}^1\Delta_g) \\ \mathrm{O}_2^*({}^1\Delta_g) + \mathrm{Car} &\rightarrow \mathrm{O}_2 + {}^3\,\mathrm{Car}^*. \end{aligned} \tag{7.3}$$

Thus, the carotenoids prevent the accumulation of singlet oxygen which causes photo-oxidation of the plant cells.

Some other functions of carotenoids such as stabilizing the lipid phase of the thylakoid membrane are known [2]. However, in our discussion of the carotenoid excited state spectrum, we will concentrate mainly on the light harvesting role of the carotenoids, i.e. the singlet excited states of the carotenoid molecules.

In the process of light harvesting, involving the absorption of visible light by carotenoids and the subsequent transfer of energy to the chlorophyll molecule, there is the involvement of low-lying bright as well as dark singlet excited states of the carotenoid molecule. For a long time, the low-lying spectra of carotenoids involved in the light harvesting process was known to consist of three states, the ground state $1A_g^-$ (known as the S_0 state), the optically allowed excited state $1B_u^+$ (known as the S_2 state) and the intermediate dark excited state $2A_g^-$ (known as the S_1 state). This three state light harvesting mechanism [3, 4] can be understood pictorially as shown in Fig. 7.1. Here it should be noted that the labels A and B in the symmetry of the ground and excited states in carotene are approximate labels (since carotenoids lack the σ_h plane of symmetry). However, this notation of labeling the excited states is used to preserve the analogy with the polyene molecules which form the backbone of the carotenoid chromophores.

7.2 Low-lying excited states of carotenoids

There are over 600 carotenoids that have been isolated to date. The common feature among all of them is the polyene backbone structure. There are some special features such as the presence of carbonyl or hydroxy groups in some of the carotenoids, e.g. peridinin and spheroidene. The effect of presence of

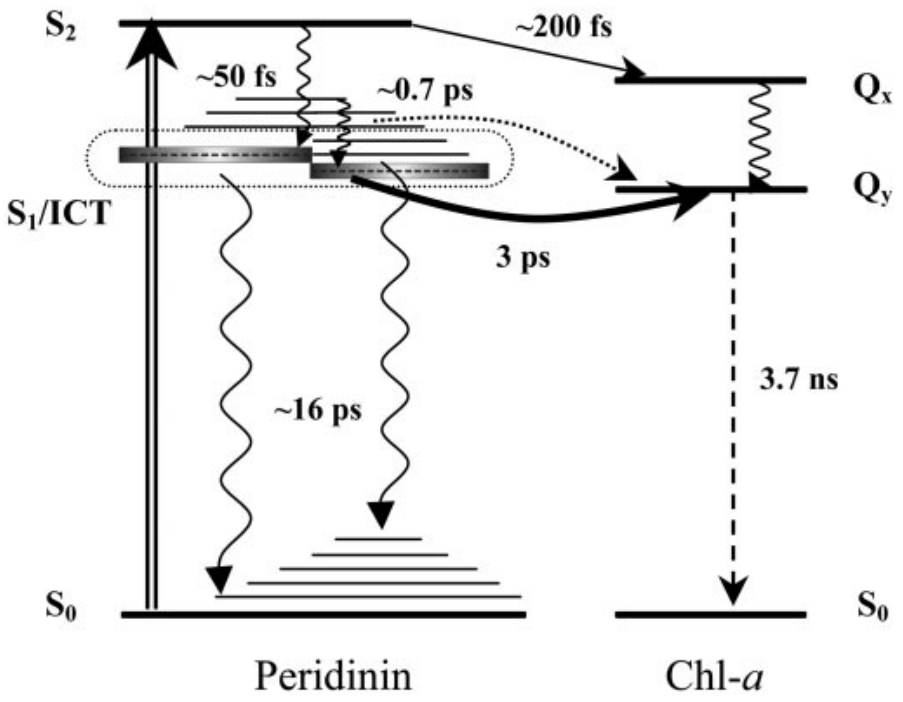


Figure 7.1: Light harvesting mechanism involving three states of carotenoids: S_0, S_1 and S_2. This was the accepted mechanism for a long time, till new dark states were observed in between S_1 and S_2 states in the carotenoids, thus complicating the scenario.

carbonyl groups in the carotenoids has been studied over the last few years [5]. However, we are going to consider only the simple hydrocarbon carotenoids in this chapter.

The relatively simple picture of three state light harvesting presented above has been questioned due to recent observation of several excited states between the S_1 and S_2 states [6, 7]. In 1999 Sashima *et al*. [8, 9] observed the first of these intermediate states in spheroidene by resonance Raman spectroscopy and gave the label S_x to this excited state. His work suggested that the symmetry of this state could be $1B_u^-$. In 2002, Cogdell and coworkers [10] observed another state S^* in β-carotene and lycopene. They measured a very fast (50 fs) relaxation of S_2

to S^* thus, casting suspicion upon the original (3 state) mechanism of light harvesting. Furuichi *et al*. [11] observed another low-lying excited state the same year in bacterial carotenoids and suggested that the state might be $3A_g^-$ state by comparison with semi-empirical results by Tavan and Schulten [12, 13]. Larsen *et al*. [14] observed $S^\ddagger$ state in 2003 in β-carotene using dispersed multipulse transient absorption. Several experiments have been carried out since then to understand the dynamics of processes during light harvesting [15].

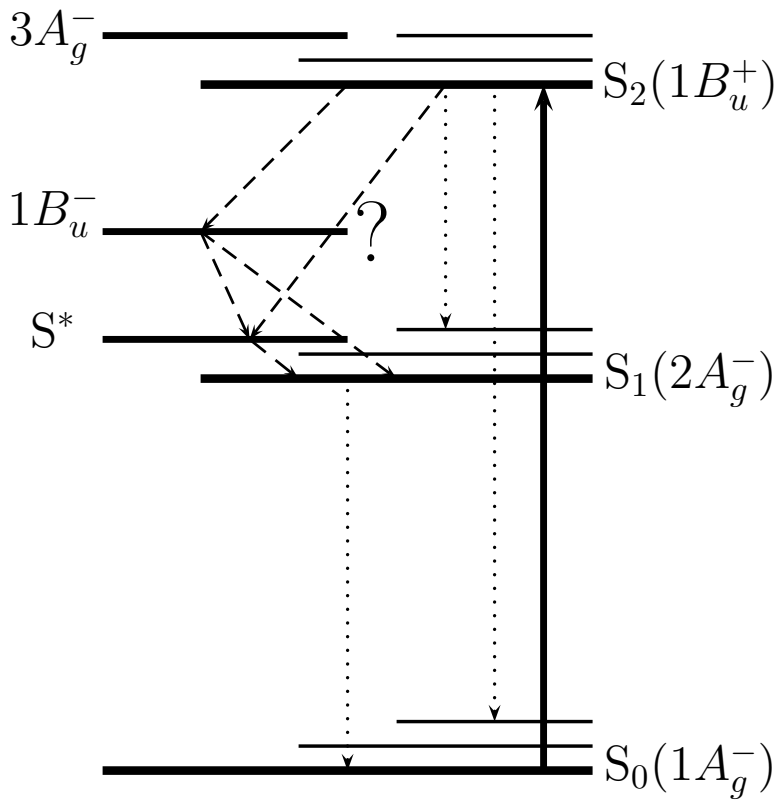


Figure 7.2: Schematic diagram of the possible mechanisms of light harvesting. With the observation of new dark states in carotenoids between S_1 and S_2, there can be several pathways through which the light harvesting process can occur.

Theoretically, the light-harvesting mechanism in carotenoids have been stud-

ied mainly by comparing with the polyene (of equivalent conjugation length) predictions made by the Parr-Pariser-Pople model Hamiltonian [12, 13]. Incomplete valence CASSCF calculations have been carried out on the *ab-initio* Hamiltonian of the rhodopin glucoside (carotenoid present in photosynthetic purple bacteria) molecule in light harvesting complex II by Sundström and coworkers [16]. Density functional theory [17–20] has proved to be a useful tool to study such systems. However, due to the importance of the multireference nature of these excited states, density functional theory has its shortcomings (explained in Chapter I). Therefore, the correct strategy to understand these excited states would be to explicitly correlate all the valence π orbitals of the hydrocarbon backbone.

The main problems that we are trying to address are:

1. The exact number and positions of the dark states between the S_1 and S_2 states in the carotenoid molecules.
2. A complete understanding of the dark states, their nature, symmetry and detailed excitonic structure.

The first step towards answering these questions would be to carry out a complete explicit electron correlation study of the valence π space. We have already tackled this part of the problem with our DMRG-CASSCF method in the Chapter IV of this thesis. However, what remains to be done is the complete analysis of all the dark states below the $S_2(1B_u^+)$ state, i.e. their excitonic nature. Since we have self-consistently solved for the correct active space and correlated the complete π space of the chromophore, we presumably have the correct qualitative nature of the excited state wavefunctions. There is also the question of the

exact number and position of the dark states, the answer to which necessitates the calculation of the exact dynamic correlation on these various excited states.

Since we already have the DMRG-CASSCF qualitative wavefunction of the excited states, we are in a position to tackle the second question.

7.3 Nature of states

Since the carotenoid backbone is very similar to that of the conjugated hydrocarbons, with minor differences due to the steric and polar effects of the ligands attached to them, we can try to understand the nature of the low-lying excited states in carotenoids in a manner analogous to the excitonic states of the linear polyenes [21].

7.3.1 Theory

An *exciton* consists of an electron and a hole bound together by Coulomb interaction. The concept of excitonic states have originally been used to describe semi-conductors in solid state physics. The idea of excitons in linear polyenes follows directly from the same concept used in solid state physics. In linear polyenes, when an electron is excited (generally by photo-excitation) from the "valence band" (formed by the occupied electronic states in the ground state) to the "conduction band" (the virtual electronic states), a positively charged hole is created in the valence band and a negatively charged electron is moved to the conduction band. The electron-hole pair are bound to each other due to Coulomb interaction and create a bound state called the "exciton".

To understand the different types of excitons that are created, we first try to understand the systems in the weakly and strongly coupled limits (which is similar to the weakly and strongly correlated systems explained in Chapter I). The HOMO-LUMO gap in weakly coupled systems is greater than the on-site Coulomb repulsion, while in strongly coupled systems the HOMO-LUMO gap is lower than the on-site Coulomb repulsion. In the intermediate coupling regime the HOMO-LUMO gap and the on-site Coulomb repulsion are similar in magnitude.

Weakly correlated systems give rise to Mott-Wannier (MW) excitons. (Actually in these systems there can be 2 kinds of excitons: *Mott-Wannier* excitons, which have long electron-hole correlation length and *Frenkel* excitons where the electron and hole are in the same position. But, for simplicity in the case of linear polyenes in the weakly coupled limit all the excitons are referred to as Mott-Wannier excitons.) Mott-Wannier excitons can be simply viewed, in the real space picture as an electron in a conduction band of localized molecular orbitals bound to a hole in the valence band molecular orbitals. The excitation operator that creates this type of excitons (we will call this the MW operator) can be written as,

$$\hat{S}_{ij}^{\dagger} = \frac{1}{\sqrt{(2)}}(\hat{a}_{i\sigma}^{c\dagger}\hat{a}_{j\sigma}^{v} \pm \hat{a}_{i\sigma'}^{c\dagger}\hat{a}_{j\sigma'}^{v\dagger}) \tag{7.4}$$

where + creates a singlet exciton and - creates a triplet exciton, and σ and σ' denote particles of opposite spin. S_{ij} denotes a particle-hole excitation from the ground state with the hole in the localized ith molecular orbital and the electron in the jth molecular orbital. The operators $\hat{a}^{c\dagger}$ creates a particle (electron) in the conduction band and $\hat{a}^{v}$ annihilates a particle (i.e. creates a hole) in the valence band. Pictorially this can be seen as the MO diagram in Fig. 7.3. In order to define a particle-hole correlation function or Mott-Wannier correlation function,

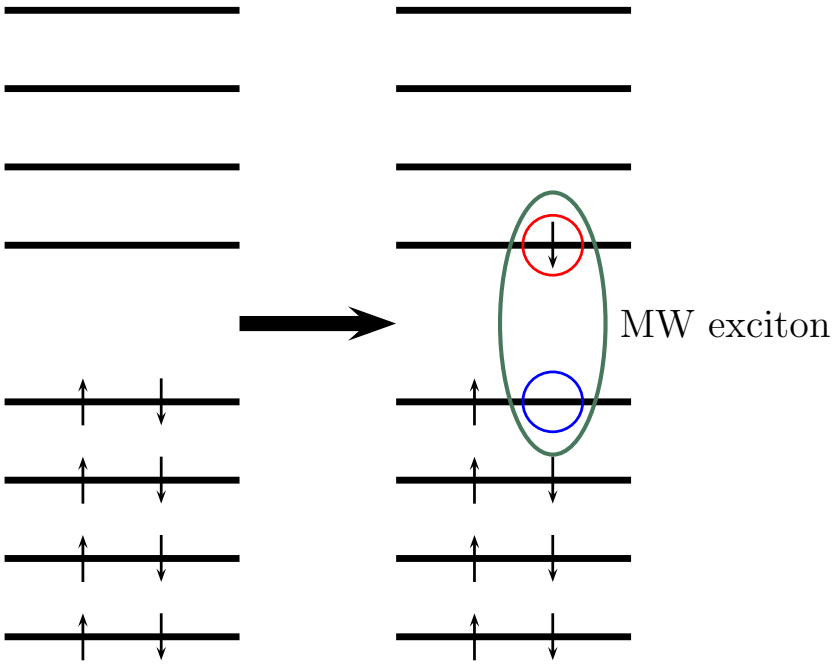


Figure 7.3: Molecular orbital picture of a Mott-Wannier exciton. The hole in the conduction band is bound to the electron excited to the valence band, thus forming the MW exciton (denoted by the dark green ellipse in the figure).

we can project the MW operator $\hat{S}_{ij}$ onto the ground state wavefunction, and obtain the dot product with the excited state wavefunction and this projected state, given by

$$\Phi(\vec{R}, \vec{s}) = \langle \Phi^{MW} | \hat{S}_{ij} | GS \rangle, \tag{7.5}$$

where $\vec{R} = \frac{\vec{r_1}+\vec{r_2}}{2}$ and $\vec{s} = \vec{r_1} - \vec{r_2}$, $\vec{r_1}$ and $\vec{r_2}$ being the position vectors of the electron and hole respectively.

However, in the strong coupling limit the situation is quite different. Instead of fully-occupied and vacant molecular orbitals, there are half-filled Hubbard orbitals, corresponding to one electron per π orbital. The excitons, thus created

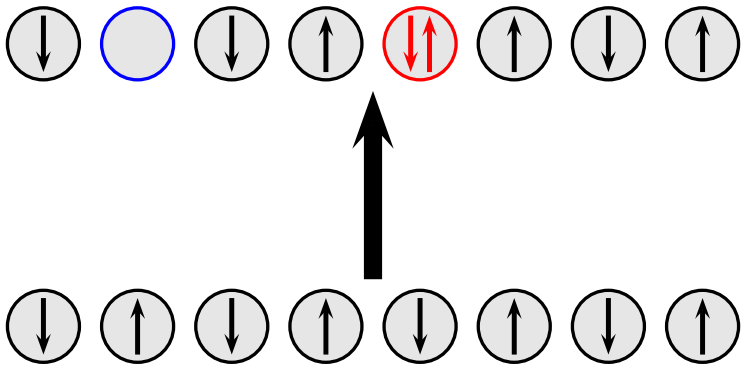


Figure 7.4: Molecular orbital picture of a Mott-Hubbard exciton. The hole in the lower Hubbard band is bound to the electron excited to the upper Hubbard band, thus forming the MH exciton.

can be of two types:

1. Holon-doublon exciton

   If one electron in the lower Hubbard band is excited to the upper Hubbard band, the hole thus created in the lower Hubbard band is bound to the electron in the upper Hubbard band. This is a *Mott-Hubbard* exciton or holon-doublon exciton. In the real space picture, an electron from one of the half filled molecular orbitals is excited into one of the other molecular orbitals, thus creating one vacant molecular orbital (hole or holon) and one doubly occupied molecular orbital (doublon) (see Fig. 7.4). A correlation function similar to that for Mott-Wannier excitons can be formulated by noting that the operators $h_i^\dagger$ and $d_i^\dagger$ defined by,

$$h_i^\dagger = \sum_\sigma a_{i\sigma}(1 - n_{i\sigma'}) \tag{7.6}$$

$$d_i^\dagger = \sum_\sigma a_{i\sigma}^\dagger n_{i\sigma'} \tag{7.7}$$

create a holon and a doublon respectively, in second quantized notation. In the Eqns. 7.6 and 7.7, σ and σ' denote particles of opposite spins and n_i denotes the occupancy of the ith orbital given by $a_i^\dagger a_i$. The MH correlation function (or holon-doublon correlation function) can be defined by the transition expectation value of the operator, $h_i^\dagger d_j^\dagger$ with the ground and excited states, given by

$$\Phi(\vec{R}, \vec{s}) = \langle \Phi^{MH} | h_i^\dagger d_j^\dagger | GS \rangle, \tag{7.8}$$

where $\vec{R} = \frac{\vec{r_1}+\vec{r_2}}{2}$ and $\vec{s} = \vec{r_1} - \vec{r_2}$, $\vec{r_1}$ and $\vec{r_2}$ are the position vectors of the holon and doublon thus created.

2. Bimagnon exciton

   When two electrons in the Hubbard band simply change their spins, the pair of magnons (singlet-triplet excitations) recouple to form a resultant *bimagnon* exciton (singlet). The pictorial description of this state is given in Fig. 7.5.

   The bimagnon operator can be written in second quantization as,

$$M_{ij} = a_{i\sigma}^\dagger a_{j\sigma'}^\dagger a_{i\sigma'} a_{j\sigma} + a_{i\sigma'}^\dagger a_{j\sigma}^\dagger a_{i\sigma} a_{j\sigma'} \tag{7.9}$$

   (a and $a^\dagger$, σ and σ' are defined as above) and thus analogous to the MW and MH cases, the bimagnon exciton correlation function can be defined as

$$\Phi(\vec{R}, \vec{s}) = \langle \Phi^{BM} | M_{ij} | GS \rangle, \tag{7.10}$$

   where $\vec{R} = \frac{\vec{r_1}+\vec{r_2}}{2}$ and $\vec{s} = \vec{r_1} - \vec{r_2}$, $\vec{r_1}$ and $\vec{r_2}$ are the position vectors of the magnons (or spin flips).

The conjugated hydrocarbons lie in the intermediate region (between the strongly and weakly coupled, i.e the HOMO-LUMO gap is comparable to the

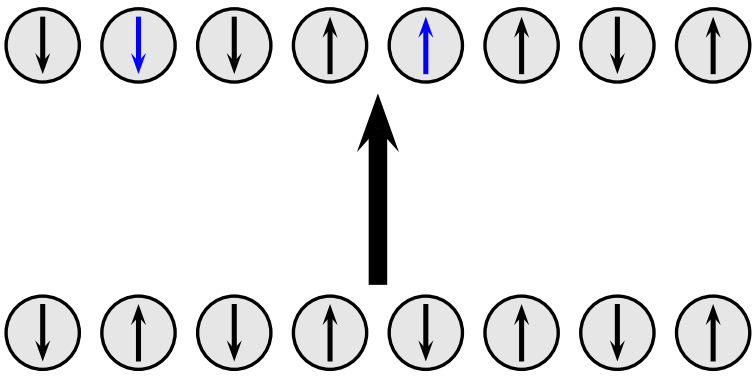


Figure 7.5: Molecular orbital picture of a bimagnon exciton. Two coupled singlet-triplet (magnon) excitations create the bimagnon exciton.

on-site Coulomb repulsion) and therefore the carotenoid molecules formed by the conjugated hydrocarbon chromophore lie in the intermediate region. Thus, the excited state multireference wavefunctions of these systems are a mixed description of *all* these excitons.

7.3.2 Computational details

The geometry of the *s-cis* β-carotene molecule was optimized in the 6-31G basis set at the density functional level using the B3LYP functional [22, 23] as implemented in GAUSSIAN03 [24]. The initial integrals were obtained from RHF calculations in PSI3 [25, 26] followed by the PAO technique described in Chapter IV and localized by Pipek-Mezey localization scheme [27].

State averaged (four state) DMRG-CASSCF calculations with $M = 250$ was performed on this molecule with intermediate localization of the active orbitals by

the PML scheme. After the active space orbitals were optimized for these four states, nine state averaged DMRG calculations were done with $M = 600$ and $M = 1000$ (M being the no. of states retained in the DMRG calculation as explained in Chapter II).

The validity of the four state averaged orbitals being used for the complete nine state averaged DMRG calculations was checked by varying the state averaging at the CASSCF level in $C_{12}H_{14}$ molecule.

7.3.3 Results and Discussion

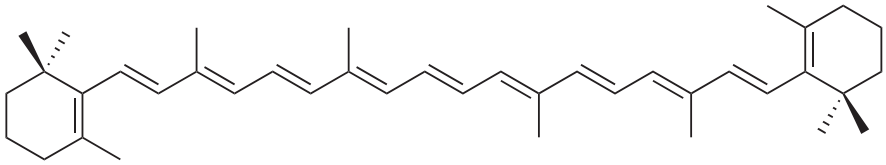


Figure 7.6: *s-cis* β-carotene.

Effect of state averaging

From table 7.1, it is clear that the excitation energies of the excited states of $C_{12}H_{14}$ (C and H atoms in Dunning's cc-pVDZ basis set [28]) are not very sensitive to the degree of state averaging in the optimization of the active space, i.e. the composition of the active space is not sensitive to the degree of state averaging. This proves the validity of our calculations on β-carotene, where we will be performing DMRG-CASSCF calculation with four state averaging and with this four state optimized active space, we will perform DMRG (equivalent to CASCI) to obtain all the nine states that are needed to target the $1B_u^+$ optically active state.

Table 7.1: Effect of orbital optimization using different state averaging schemes. n state averaged refers to orbital optimization with n state averaged CASSCF followed by 5 state averaged CASCI calculation to obtain the excitation energies of the low-lying excited states of $C_{12}H_{14}$ (in eV). The basis set used is Dunning's cc-pVDZ.

State	Excitation energy		
	2 state averaged	3 state averaged	4 state averaged
$2A_g^-$	3.84	3.79	3.77
$1B_u^-$	4.86	4.77	4.74
$3A_g^-$	5.75	5.66	5.60
$1B_u^+$	6.03	6.00	5.99

Position of excited states

Table 7.2 shows the excitation energies and oscillator strengths of the *s-cis* β-carotene molecule excited states obtained by DMRG-CASSCF calculations with 6-31G basis set. The excitation energies are over-estimated mainly due to the absence of dynamic correlation in our calculations. From other theoretical work on long polyenes [29], it follows that carotene (and long polyene) excited states (especially the optically active $1B_u^+$ state) are considerably influenced by dynamic correlation. Another possible cause for the discrepancy between the experimental and calculated excitation energies can be from the geometric relaxation of the excited states.

Thus, while our DMRG-CASSCF calculation excitation energies are not quantitatively correct, they do show that there are a number of dark states that are below or near the $1B_u^+$ state and the correct number and position of dark states

Table 7.2: The excitation energies (in eV) and oscillator strengths (in $a.u.$) of the β-carotene excited states calculated by using 4 state averaged DMRG-CASSCF ($M = 250$) followed by 9 state averaged DMRG ($M = 1000$).

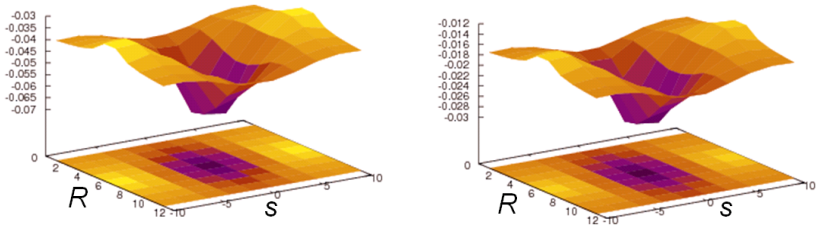
State	Excitation Energy	Osc. Strength	Expt.
$2A_g^-$	2.99	Forbidden	1.76 [1]
$1B_u^-$	3.63	0.2025	2.18 [2]
$3A_g^-$	4.32	Forbidden	2.22 [2]
$4A_g^-$	4.55	Forbidden	-
$2B_u^-$	4.91	0.1693	-
$3B_u^-$	5.20	0.1930	-
$5A_g^-$	5.39	Forbidden	-
$1B_u^+$	5.50	0.7824	2.47 [2]

[1][9] (for lycopene and β-carotene).
[2]Excitation measured for lycopene [11] (for lycopene).

can only be predicted after the consideration of the dynamic correlation in this problem. However, since the static correlation is generally sufficient to obtain the correct qualitative description of the excited states, we are now in a position to analyze and understand the nature of these dark states.

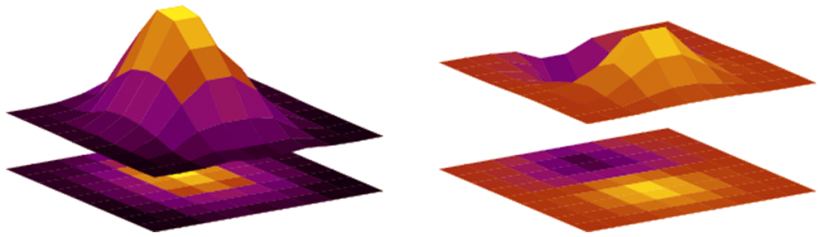
Nature of excited states

From the definitions of the Mott-Wannier, Mott-Hubbard and bimagnon correlation functions, we can calculate the relative importance of these kinds of excitons to the low-lying excited states in β-carotene. The $\%$ weight of these different excitons in the excited states are listed in the Table 7.3. We can further obtain the quantum numbers associated with the center of mass and relative



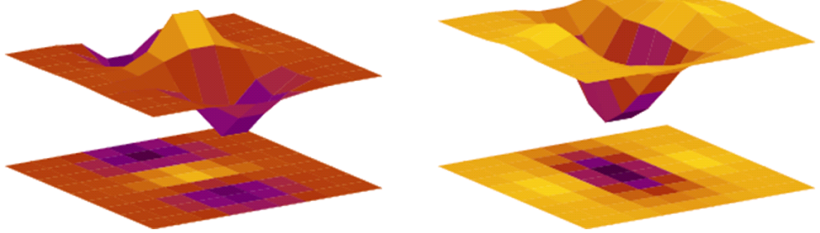
(a) Mott-Wannier exciton. (b) Mott-Hubbard exciton.

Figure 7.7: Mott-Wannier and Mott-Hubbard correlation functions corresponding to $1B_u^+$ state (see Eqns. 7.5 and 7.8). There are no nodes in the correlation function when plotted against $\vec{R} = \vec{r_1} + \vec{r_2}$ and $\vec{s} = \vec{r_1} - \vec{r_2}$, $\vec{r_1}$ and $\vec{r_2}$ being the position of the electron and hole (in case of MW exciton) and holon and doublon (in case of MH exciton) in units of repeat distance.



(a) $2A_g^-$ state is made up of bimagnon exciton $n = 1$, $K = 1$.

(b) $1B_u^-$ state is made up of bimagnon exciton $n = 1$, $K = 2$.



(c) $3A_g^-$ state is made up of bimagnon exciton $n = 1$, $K = 2$.

(d) $4A_g^-$ state is made up of bimagnon exciton $n = 2$, $K = 1$.

Figure 7.8: Bimagnon correlation function corresponding to $2A_g^-$, $1B_u^-$, $3A_g^-$ and $4A_g^-$ states (see Eqn. 7.10). The number and position of nodes reflect the n and k values listed in the table 7.3.

Table 7.3: Nature of the β-carotene excited states. $\%$ of the Mott-Wannier (MW), Mott-Hubbard (MH) and Bimagnon (BM) nature of excited states are tabulated. The n and K values denote the center of mass momemtum quantum number and the quantum number associated with the reversal of relative coordinates. The values of n and K reported are for the type of exciton that forms the majority of the contribution to the respective excited states.

State	MW nature	MH nature	BM nature	n	K
$2A_g^-$	8	1	91	1	1
$1B_u^-$	10	1	89	1	2
$3A_g^-$	12	1	87	1	3
$4A_g^-$	6	1	93	2	1
$2B_u^-$	13	9	86	1	4
$3B_u^-$	8	1	91	1	5
$5A_g^-$	14	1	85	1	6
$1B_u^+$	81	17	2	1	1

coordinate momenta, corresponding to the excitons, by plotting the correlation functions in the real and conjugate space and noting the number and position of nodes in the correlation functions(see Figs. 7.7 and 7.8).

From the table 7.3, we observe that all the optically forbidden B_u^- and A_g^- excited states below the $1B_u^+$ state is made up of bimagnon excitons, i.e. in the valence bond language they are formed by the coupling of two singlet to triplet excitations. On the other hand, the optically allowed state $1B_u^+$ can be described quite well as a Mott-Wannier exciton, i.e. described well within a single determinant weakly interacting picture. This re-asserts our knowledge about the

$1B_u^+$ state that it is made predominantly by the HOMO-LUMO transition and the dynamic correlation forms the most important part of the electronic correlation needed to describe this state. (Due to the predominant HOMO-LUMO transition and single-configurational character, static correlation plays a minor role in this state.)

7.4 Excitation energies - Inclusion of dynamic correlation

There is yet another major question that remains to be answered in the carotenoid excitation spectra: the precise position of the dark states relative to the $1B_u^+$ state, which may be involved in the process of light harvesting. As explained in the previous section, static correlation is not enough to answer this question. We need to include dynamic correlation on top of the static correlated qualitatively correct wavefunctions to obtain quantitatively correct excitation energies.

In the Chapter V and VI, we have described two dynamic correlation methods that can be applicable to large strongly correlated systems: cumulant approximated NEVPT2 and canonical transformation theory.

7.4.1 Cumulant approximated NEVPT2

Cumulant approximated NEVPT2 (cu-SC-NEVPT2) is the relatively computationally cheaper and easier way of incorporating dynamic correlation in the strongly correlated systems, although not the most accurate.

Computational details

The geometry of lycopene molecule ($C_{40}H_{56}$) was optimized by the same procedure as used for β-carotene. The C_{22} backbone (almost planar which forms the most important component of the chromophore) coordinates of the optimized lycopene molecule were taken, and H atoms were attached to this C_{22} backbone, followed by the re-optimization of the H atoms keeping the C_{22} framework fixed at the density functional level using the B3LYP functional [22, 23] as implemented in GAUSSIAN03 [24] using 6-31G basis set. The initial localized orbitals for the "model lycopene" molecule (out of plane $C_{22}H_{24}$) with Dunning's cc-pVDZ basis [28] for the C atoms and STO-6G basis for the H atoms was obtained from PSI3 [25, 26].

The DMRG-CASSCF (4 state averaged with $M = 250$) and subsequent DMRG (9 state averaged with $M = 1000$) calculations were performed exactly as in the case of β-carotene. The information about the one and two particle density matrices along with the final optimized wavefunction was used to calculate the cumulant approximated NEVPT2 (cu-SC-NEVPT2) energies.

Results

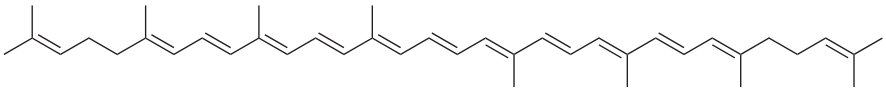


Figure 7.9: Lycopene.

Table 7.4 shows the excitation energies in eV of the six lowest energy states (in the cu-SC-NEVPT2 level) of lycopene (model compound). Although the $2A_g^-$, $1B_u^-$ and $3A_g^-$ excitation energies matched reasonably well with the experimentally observed excitation energies, the ordering of the ground and excited states

Table 7.4: Excitation energies (in eV) of the lycopene model compound by DMRG-CASSCF and cu-SC-NEVPT2 methods. The basis and geometry as explained in text.

State	DMRG-CASSCF	cu-SC-NEVPT2	Expt.
$2A_g^-$	2.80	1.67	1.76 [1]
$1B_u^-$	3.36	2.71	2.18 [1]
$3A_g^-$	3.95	2.80	2.22 [2]
$4A_g^-$	4.15	3.30	-
$2B_u^-$	4.51	5.25	-
$1B_u^+$	5.16	1.21	2.47 [2]

[1][9] (for lycopene and β-carotene).
[2]Excitation measured for lycopene [11] (for lycopene).

was found to be $1A_g^- < 1B_u^+ < 2A_g^- < 1B_u^- < 3A_g^- < 2B_u^- < 4A_g^-$. The positioning of $1B_u^+$ state below the $2A_g^-$ state is almost certainly an artifact of the calculation. Since the cumulant approximated NEVPT2 is not a very powerful dynamic correlation method, it seems to have problems describing states which have a higher degree of dynamic correlation (e.g. $1B_u^+$).

There is one more concern about using this method to obtain the correct excitation energy and number of the dark states. As explained in Chapter V, this method uses an imaginary level shift to correct for possible false intruder states due to the cumulant errors in the higher body reduced density matrices. In the case of carotenoids, the energy of the ground and excited states are very sensitive to the value of level shift used and level shift analysis is inadequate to decide on an appropriate level shift. The values listed in the table 7.4 are with no level shift.

7.4.2 CT theory

Due to the uncertainty of the level shift analysis in cumulant approximated NEVPT2, a more rigorous method for calculating dynamic correlation would be to calculate the strongly contracted CT energies (see Chapter VI) with the state averaged DMRG-CASSCF reference wavefunction and the one- and two-particle density matrices of the individual states.

7.5 Conclusion and future directions

In this chapter, we have described the application of DMRG-CASSCF and dynamic correlation methods like cu-NEVPT2 and CT theory to understand the nature and positions of the low-lying excitations in carotenoids, like β-carotene and lycopene. While our DMRG-CASSCF calculations get the correct qualitative picture of these low-lying dark and optically active states, we are yet to obtain sufficient quantitative accuracy in our calculations. Extensively state averaged DMRG-CASSCF calculations on these compounds give us the excitonic nature of the low-lying excitations.

However, the dynamic correlation methods are still not adequate to handle such large and complicated systems. The cu-SC-NEVPT2 calculations give a reasonable estimate of the dynamic correlation in the dark states (which are not extremely dynamically correlated). However, the predominantly dynamically correlated $1B_u^+$ state excitation energy obtained by this method is not accurate as the perturbative technique (in the approximate form) is not powerful enough to handle the high degree of dynamic correlation. Thus, the addition of dynamic

correlation with sufficient accuracy requires further work.

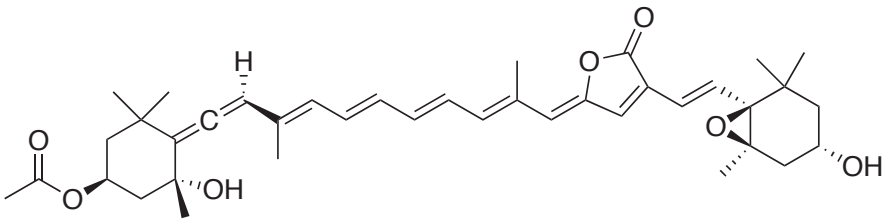


Figure 7.10: Peridinin.

There is another aspect of the carotenoid problem that requires attention: the effect of conjugated carbonyl groups to the excitation spectrum in carotenoids, like peridinin. With our DMRG-CASSCF, we have a technique that is capable of explicitly correlating the complete π valence space. Therefore, we have a tool to understand the nature of the excited state in the carotenoid molecules containing carbonyl and hydroxy groups. This would be an interesting direction for future work.

BIBLIOGRAPHY

[1] H. A. Frank, Archives of Biochemistry and Biophysics **385**, 53 (2001).

[2] B. Ke, *Photosynthesis: photobiochemistry and photobiophysics*, volume 10 of *Advances in Photosynthesis*, chapter Role of Carotenoids in Photosynthesis, page 229, Kluwer Academic Publishers, Boston, 2001.

[3] V. Sundström, Progress in Quantum Electronics **24**, 187 (2000).

[4] D. Zigmantas, R. G. Hiller, V. Sundström, and T. Polivka, Proc. Nat. Acad. Sci. **99**, 16760 (2002).

[5] D. Zigmantas et al., Phys. Chem. Chem. Phys. **6**, 3009 (2004).

[6] R. Fujii et al., J. Phys. Chem. A **105**, 5348 (2001).

[7] K. Onaka et al., Chem. Phys. Lett. **315**, 75 (1999).

[8] T. Sashima, H. Nagae, M. Kuki, and Y. Koyama, Chem. Phys. Lett. **299**, 187 (1999).

[9] T. Sashima, Y. Koyama, T. Yamada, and H. Hashimoto, J. Phys. Chem. B **104**, 5011 (2000).

[10] G. Cerullo et al., Science **298**, 2395 (2002).

[11] K. Furuichi, T. Sashima, and Y. Koyama, Chem. Phys. Lett. **356**, 547 (2002).

[12] P. Tavan and K. Schulten, J. Chem. Phys. **85**, 6602 (1986).

[13] P. Tavan and K. Schulten, Phys. Rev. B **36**, 4337 (1987).

[14] D. S. Larsen et al., Chem. Phys. Lett. **381**, 733 (2003).

[15] P. J. Walla, P. A. Linden, C.-P. Hsu, G. D. Scholes, and G. R. Fleming, Proc. Nat. Acad. Sci. **97**, 10808 (2000).

[16] Z. He, V. Sundström, and T. Pullerits, Chem. Phys. Lett. **334**, 159 (2001).

[17] C.-P. Hsu, P. J. Walla, M. Head-Gordon, and G. R. Fleming, J. Phys. Chem. B **105**, 11016 (2001).

[18] A. Dreuw, G. R. Fleming, and M. Head-Gordon, Phys. Chem. Chem. Phys. **5**, 3247 (2003).

[19] S. Schlücker, A. Szeghalmi, M. Schmitt, J. Popp, and W. Kiefer, Journal of Raman Spectroscopy **34**, 413 (2003).

[20] K. V. Berezin and V. V. Nechaev, Journal of Applied Spectroscopy **72**, 164 (2005).

[21] W. Barford, *Electronic and Optical Properties of Conjugated Polymers*, Oxford University Press, 1999.

[22] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[23] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[24] M. J. Frisch et al., GAUSSIAN 03, REVISION C.02, Gaussian, Inc., Wallingford CT, 2004, see http://www.gaussian.com/.

[25] T. D. Crawford et al., PSI 3.2 (2003), see www.psicode.org.

[26] T. D. Crawford et al., J. Comput. Chem. **28**, 1610 (2007).

[27] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

[28] T. H. Dunning Jr., J. Chem.Phys. **90**, 1007 (1989).

[29] Y. Kurashige, H. Nakano, Y.Nakao, and K. Hirao, Chem. Phys. Lett. **400**, 425 (2004).

CHAPTER 8

FUTURE DIRECTIONS: IMPROVING THE STATIC AND DYNAMIC CORRELATION METHODS

8.1 Need for improvement in the static correlation methods

DMRG and its orbital optimized counterpart give us a method to accurately obtain the static correlation in strongly correlated "pseudo one dimensional" systems of essentially infinite length. However, the method needs to be extended to treat general large systems in more than one dimension.

In order to establish a viable strategy to extend the DMRG algorithm to higher dimensions let us consider a general two dimensional lattice and try to use DMRG to solve for the static correlation in the lattice. The lattice sites must be ordered in one dimension and one possible ordering scheme can be as shown in Fig. 8.1. However, as can be seen from the figure this ordering scheme does not account for several of the required correlations, those between the sites on the rows. The current strategy in such situations is to retain a larger number of states M in the DMRG calculation so that, even with the absence of direct correlations between these sites, we can recover some of the correlations indirectly through the large number of states that are retained. Thus, with the increase in size of the system in its second dimension, M needs to be increased and many of the natural benefits of the DMRG algorithm in one dimension are lost.

However, for a better method that can naturally handle two dimensional systems we need to incorporate these correlations without the extra computational cost of increasing M states that are retained. Following the *matrix product state*

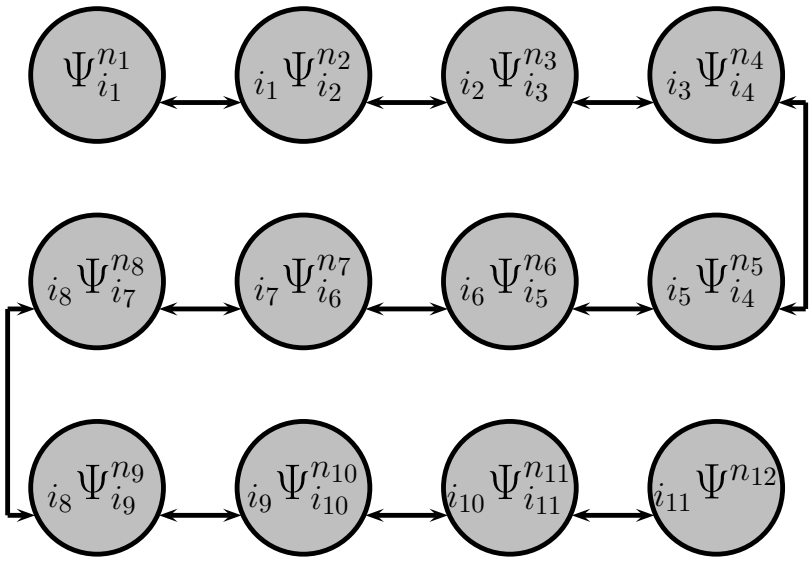


Figure 8.1: DMRG applied to a 2D lattice: as can be seen from the figure, this scheme is inefficient to consider correlations between the rows of sites.

(MPS) formulation, intuitively we can think of incorporating the correlations by using another set of auxiliary indices, j. Thus, if the MPS wavefunction is given by,

$$\Psi^{n_1 n_2 n_3 \cdots} \approx \sum_{n_1 n_2 n_3 \ldots} \sum_{i_1 i_2 i_3 \ldots} \Psi^{n_1}_{i_1} \Psi^{n_2}_{i_1 i_2} \Psi^{n_3}_{i_2 i_3} .. \tag{8.1}$$

the new wavefunction which can be viewed as the *tensor product state* can be given by,

$$\Psi^{n_1 n_2 n_3 \cdots} \approx \sum_{n_1 n_2 n_3 \ldots} \sum_{i_1 i_2 i_3 \ldots} \sum_{j_1 j_2 j_3 \ldots} \Psi^{n_1}_{i_1 j_1} \Psi^{n_2}_{i_1 i_2 j_2} \Psi^{n_3}_{i_2 i_3 j_3} \ldots \tag{8.2}$$

where the extra indices $j_1, j_2 \cdots$ introduce correlations that were originally absent in the matrix product state formulation and yet were necessary due to the second dimension in the problem.

This class of tensor product states (TPS) has been used in quantum information theory and of late been used to solve certain Bosonic problems and spin Hamil-

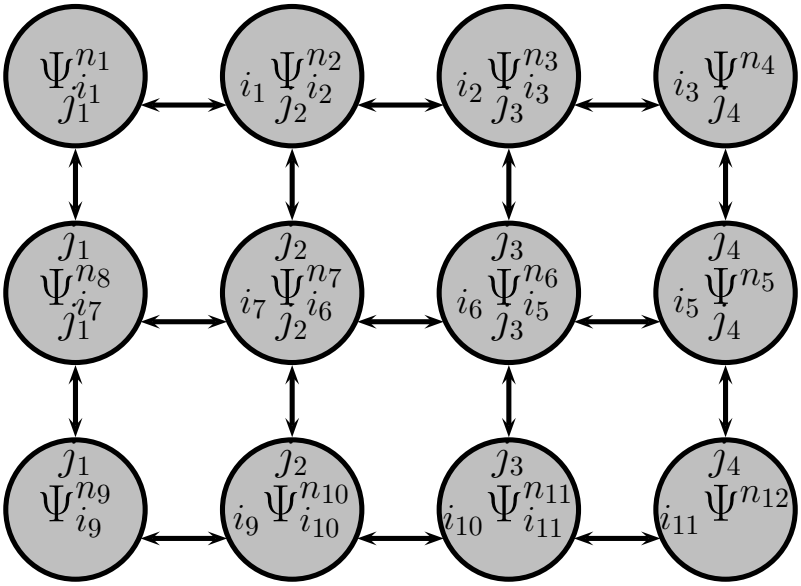


Figure 8.2: Projected entangled pair states applied to a 2D lattice: here the auxiliary index j introduces the correlation between the rows of sites.

tonians by the ansatz known as projected entangled pair states (PEPS) [1, 2] and multi-scale entanglement renormalization ansatz (MERA) [3]. Recently, there has been some efforts to use these methods for fermionic systems. However, there are generally many numerical issues with the equations and therefore, a stable and efficient implementation of these methods for quantum chemical systems is an interesting direction for further research.

8.2 Need for improvement in the dynamic correlation methods

In this thesis, we have described two dynamic correlation methods for large multireference systems: cumulant approximated NEVPT2 and canonical transformation theory. However, from the applications of the methods to large sys-

tems, we know that the methods have several short comings. There are two aspects of the dynamic correlation methods that need improvement.

1. Accuracy

   Cumulant approximated NEVPT2 has its inadequacies when treating systems that have a very large dynamic correlation while being inherently multireference (see Chapter VII). CT theory also has convergence problems with such complicated systems. One of the main causes of these problems in both these methods is the cumulant approximation of the higher (more than 2) body terms. Thus, this problem can be ameliorated by introducing some more information about the higher body reduced density matrices. While this would have some computational cost overhead, the creation of the three body reduced density matrix from the DMRG wavefunction can be possible. Therefore, using the information about the 3RDM from the DMRG wavefunctions the accuracy and convergence of both the methods (cu-SC-NEVPT2 and CT) can be improved.

2. Scalability

   The computational scaling of the CT theory, while better than the multireference CI methods, is quite steep ($O(n_{act}^2 n_{vir}^4) \approx O(n^6)$ where n is the system size or total number of orbitals, n_{act} and n_{vir} are the numbers of active orbitals and virtual orbitals respectively). Therefore, any attempt to use CT theory for very large systems would become prohibitively expensive. For a method to be truly applicable for very large systems the scaling should be linear or quadratic with the system size. Since electron correlations are essentially short-range, it also seems physical that the computational scaling of most methods should not be higher than quadratic

scaling. However, the reason for the unphysically high scaling of most methods is the canonical nature of the orbitals which forces us to correlate all the orbitals while in reality two electrons that are sufficiently far away from each other hardly interact. Therefore, by using localized orbitals and constraining the excitations to contain only the important ones (i.e. the excitations from and to orbitals that are adjacent or near to each other), we can achieve a local CT theory.

8.3 Evaluation of three-particle reduced density matrix from the DMRG wavefunction

The elements of the three-particle reduced density matrix (3RDM) are given by,

$$\gamma_{ijklmn} = \langle a_i^\dagger a_j^\dagger a_k^\dagger a_l a_m a_n \rangle. \tag{8.3}$$

Since the number of elements in the 3RDM is k^6, where k is the number of the lattice sites, naively one would think that the evaluation of the 3RDM would scale as $O(k^6)$ with a storage cost $O(M^2k^6)$ since each of the terms has a $M \times M$ matrix representation. The computational as well as storage cost would be prohibitively expensive. However, we can use the canonical representations at different lattice sites similar to that used for the calculation of lower body reduced density matrices (in Chapter III) and reduce the cost of the three-particle reduced density matrix calculation.

The pseudo-code for the three-particle RDM evaluation is given in Algs. (3) and (4). Alg. (4) describes how to partition the evaluation of different density matrix elements amongst the block configurations as we traverse a DMRG sweep.

Algorithm 3: COMPUTE$(nl, np, nr, left, sitep, right)$. Note $nl, np, nr \leq 3$ and $nl + np + nr = 6$, i.e. the number of indices in the three-particle density matrix γ.

for all opl= operators with nl indices on block $left$ **do**

(If parallel, loop only over opl stored on current proc)

for all opp= operators with np indices on block $sitep$ **do**

for all opr= operators with nr indices on block $right$ **do**

$$\gamma(np, nl, nr) = parity(opl, opp, opr) \times \langle \Psi | opl \otimes opp \otimes opr | \Psi \rangle$$

end for

end for

end for

(If parallel, accumulate contributions from all procs to root processor)

The different functions that are put together in Alg. (4) to form the complete three-particle density matrix can be understood by noting that there are a few major types of density matrix elements depending on how many operators are situated on different lattice sites.

- γ_{ijklmn} such that 4 operators are on one site which gives rise to COMPUTE functions with arguments $[4, 0, 2]$, $[4, 2, 0]$, $[4, 1, 1]$ in the first configuration and $[0, 4, 2]$, $[2, 4, 0]$, $[1, 4, 1]$ in the sweep through the blocks and $[0, 2, 4]$, $[2, 0, 4]$, $[1, 1, 4]$ in the final configuration.
- γ_{ijklmn} such that 3 operators are on one site which gives rise to COMPUTE functions with arguments $[2, 3, 1]$, $[1, 3, 2]$, $[3, 3, 0]$ in the the sweep through the blocks and $[0, 3, 3]$, $[1, 2, 3]$, $[2, 1, 3]$ in the final configuration.
- γ_{ijklmn} such that 2 operators are on one site, which gives rise to COMPUTE functions with arguments $[3, 2, 1]$ in the sweep though the blocks.

Algorithm 4: Evaluating three-particle density matrix by assembling across a DMRG sweep.

$left$= site 1, $sitep$= site 2, $right$= sites $3 \ldots k$

COMPUTE(4, 1, 1, $left$, $sitep$, $right$)

COMPUTE(4, 0, 2, $left$, $sitep$, $right$)

COMPUTE(4, 2, 0, $left$, $sitep$, $right$)

for $sitep$= 2 to k-1 **do**

  $left$= sites $1 \ldots p-1$, $right$= sites $p+1 \ldots k$

  COMPUTE(1, 4, 1, $left$, $sitep$, $right$)

  COMPUTE(0, 4, 2, $left$, $sitep$, $right$)

  COMPUTE(2, 4, 0, $left$, $sitep$, $right$)

  COMPUTE(2, 3, 1, $left$, $sitep$, $right$)

  COMPUTE(1, 3, 2, $left$, $sitep$, $right$)

  COMPUTE(3, 3, 0, $left$, $sitep$, $right$)

  COMPUTE(3, 2, 1, $left$, $sitep$, $right$)

  COMPUTE(2, 2, 2, $left$, $sitep$, $right$)

  COMPUTE(3, 1, 2, $left$, $sitep$, $right$)

end for

$left$= sites $1 \ldots k-2$, $sitep$= site $k-1$, $right$= site k

COMPUTE(1, 1, 4, $left$, $sitep$, $right$)

COMPUTE(2, 0, 4, $left$, $sitep$, $right$)

COMPUTE(0, 2, 4, $left$, $sitep$, $right$)

COMPUTE(0, 3, 3, $left$, $sitep$, $right$)

COMPUTE(1, 2, 3, $left$, $sitep$, $right$)

COMPUTE(2, 1, 3, $left$, $sitep$, $right$)

- γ_{ijklmn} such that only one operator is on one site, which gives rise to COMPUTE functions with arguments $[3, 1, 2]$ in the sweep through the blocks.

The actual calculation of the density matrix elements is carried out by the function COMPUTE in Alg. (3), which computes all density matrix elements that may be assembled from nl index operators on the left block, np index operators on site p, and nr index operators on the right block.

The resultant computational complexity can be understood if we look more closely at the density matrix element $\gamma_{ijklmn}\ \forall\ i \neq j \neq k \neq l \neq m \neq n$ since these are the most computationally expensive elements. In the pseudo-code it is calculated using the COMPUTE function with $[3, 1, 2]$ arguments (Alg. 5), i.e. the left block has 3 indices, the site (dot) has one index and the right block has 2 indices. That is we choose a block configuration such that i, j, k lie in the left block, l in the dot in between the blocks and m, n in the right block, i.e. $\boxed{\cdots \bullet_i \cdots \bullet_j \cdots \bullet_k \cdots}\ \bullet_l\ \boxed{\cdots \bullet_m \cdots \bullet_n \cdots}$. The corresponding matrix element may then be evaluated using $a_i^\dagger a_j^\dagger a_k^\dagger$ on the left block, and a_l on site and $a_m a_n$ on right block, and thus no operator matrices with more than three orbital indices appear on either block (see Fig. 8.3).

By the appropriate choice of partitioning between the left and right blocks, we can arrange things such that we never manipulate operators with more than three orbital labels on either the left or right blocks for any $ijklmn$. During a DMRG sweep we iterate through all block configurations where the dividing site $\bullet_p$ ranges from site 2 to site $k-1$. At each block configuration, we then evaluate all the three-particle density matrix elements which do not require more than three-index operators on either the left or right blocks, and assemble the contributions of all the block configurations at the end of the DMRG sweep. The

Algorithm 5: Pseudo-code for COMPUTE$(3, 1, 2)$ function called during the sweep through the block configuration in the three-particle density matrix calculation.

Initialize array of length $i \geq j \geq k \in left$

for all $i \in left$ **do**

  for all $j \leq i \in left$ **do**

    for all $k \leq j \in left$ **do**

      Build operator $\hat{o}_i\hat{o}_j\hat{o}_k$ save in array

    end for

  end for

end for

for all $iproc \in 1, 2 \ldots nproc$ **do**

  for all $l \in \bullet$ **do**

    for all $m \in right$ **do**

      for all $n \leq m \in right$ **do**

        Load operator $\hat{o}_m\hat{o}_n$

        Build operator $\hat{o}_l\hat{o}_m\hat{o}_n$

        $\gamma_{ijklmn} = parity(ijklmn)\langle\Psi|\hat{o}_i\hat{o}_j\hat{o}_k \otimes \hat{o}_l\hat{o}_m\hat{o}_n|\Psi\rangle$

      end for

    end for

  end for

end for

γ_{135689}

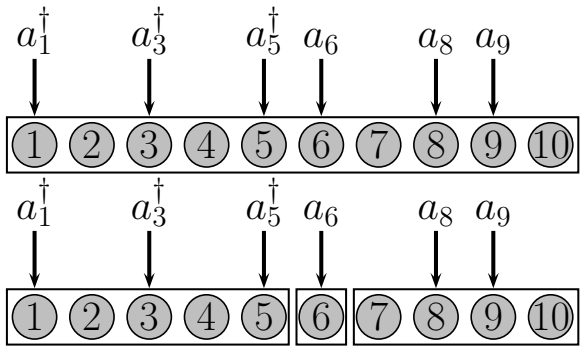


Figure 8.3: Evaluation of a 3-RDM element γ_{135689}. We can obtain this element e.g. at the block configuration where indices $1, 3, 5$ are on the left block and indices 6 on the site • and $8, 9$ are on the right block (corresponding to calling COMPUTE$(3, 1, 2)$ in Alg. 5.

total memory and storage cost is thus reduced to $O(M^2k^3)$.

We consider the different steps in building these elements.

- The creation of the operators $a_i^\dagger a_j^\dagger a_k^\dagger$ on left block and $a_l a_m a_n$ on site (in between the blocks) and right block which has k^3 matrix multiplications of size $M \times M$ thus giving rise to the computational cost of M^3k^3.
- The final multiplication of these operators on each block to get the final density matrix element $\gamma_{ijklmn} = \sum_{A,B} L_{ijk}(A, B)R_{lmn}(A, B)$ where there are k^6 indices $ijklmn$ and A and B are each of dimension M resulting to the computational cost of M^2k^6.

Therefore, the total computational cost is $O(M^2k^6 + M^3k^3)$. The computational costs of each of the different objects that build up the complete three-particle

Table 8.1: Computational cost of various steps of the evaluation of the three-particle density matrix. k is the no. of orbitals and M is the number of states retained after each decimation step in the DMRG algorithm.

Configuration	Args. to the COMPUTE functions	Total computational cost
First	$[4,1,1]$	$O(M^3k)$
	$[4,0,2]$	$O(M^3k^2)$
	$[4,2,0]$	$O(M^3)$
Sweep	$[1,4,1]$	$O(M^3k^2)$
	$[0,4,2]$	$O(M^3k^2)$
	$[2,4,0]$	$O(M^3k^2)$
	$[2,3,1]$	$O(M^3k^3)$
	$[1,3,2]$	$O(M^3k^3)$
	$[3,3,0]$	$O(M^3k^3)$
	$[3,2,1]$	$O(M^3k^3 + M^2k^6)$
	$[2,2,2]$	$O(M^3k^2 + M^2k^6)$
	$[3,1,2]$	$O(M^3k^3 + M^2k^6)$
Final	$[1,1,4]$	$O(M^3k)$
	$[2,0,4]$	$O(M^3k^2)$
	$[0,2,4]$	$O(M^3)$
	$[0,3,3]$	$O(M^3)$
	$[1,2,3]$	$O(M^3k)$
	$[2,1,3]$	$O(M^3k^2)$

reduced density matrix are given in Table 8.1.

Thus, exploiting the canonical property of the DMRG wavefunction we can formulate an efficient algorithm to evaluate the three-particle density matrix with a total per-sweep computational cost of $O(M^3k^3 + M^2k^6)$ and a memory cost of $O(M^2k^3)$.

8.4 Lower scaling canonical transformation method

CT theory [4, 5] has been introduced (in Chapter VI) as a rigorous size-extensive theory capable of describing dynamic correlation in bonding situations where there is also significant static correlation. It has also been applied to excited state problems with considerable dynamic and static correlations. Despite the very desirable properties of size-extensivity and high accuracy, there is need to reduce the computational cost scaling ($O(n_{act}^2 n_{vir}^4) \approx O(n^6)$ where n is the system size) to $O(n^2)$ in order to use it for really large systems.

Lower and affordable cost scaling is known to be attainable by exploiting the short range nature of the electron correlation. Local correlation methods such as local coupled cluster [6] and local Møller Plesset perturbation theory [7] have been developed which recover about $\approx 98\%$ of the correlation energy obtained from a corresponding non local full correlation calculation. Analogous to these methods, a local canonical transformation method can be envisaged.

It is intuitively clear that for a localized treatment of the electron correlation, we need localized orbitals. Localized fully-occupied and active orbitals can be easily obtained by separately localizing these subspaces using Pipek-Mezey localization scheme [8]. However, the virtual orbitals cannot be localized satisfactorily with this method. Orthonormal linearly independent local virtual orbitals

can be generated from projected atomic orbital (PAO) localization. A *complementary* virtual orbital space can be obtained by projecting out the core and active orbitals from the complete space of atomic orbitals. We can define a projection operator in the occupied space (containing the core or fully-occupied and the active spaces),

$$\begin{aligned} \hat{\boldsymbol{P}} &= \sum_{i=1}^{n_{occ}} |\Psi_i\rangle\langle\Psi_i| \\ &= \sum_{\mu\nu i} |\Phi_\mu\rangle C_{\mu i} C_{\nu i} \langle\Phi_\nu| \\ &= \sum_{\mu\nu} |\Phi_\mu\rangle D_{\mu\nu} \langle\Phi_\nu| \end{aligned} \tag{8.4}$$

where C denotes the molecular orbital coefficients, $D_{\mu\nu} = \sum_{i=1}^{n_{occ}} C_{\mu i} C_{\nu i}$ denotes the density matrix elements and Φ refers to atomic orbitals (AO) while Ψ refers to the molecular orbitals (MO). Thus, the projection of the AO basis on the occupied space is given by,

$$\hat{\boldsymbol{P}}|\Phi_\mu\rangle = \sum_\nu |\Phi_\nu\rangle (P^{AO}\boldsymbol{S})\mu\nu = \sum_\nu |\Phi_\nu\rangle \boldsymbol{P}_{\mu\nu}. \tag{8.5}$$

where $P^{AO}_{\mu\nu} = \sum_{i\in occ} C_{\mu i} C_{\nu i}$ and $\boldsymbol{S}$ is the overlap matrix in the AO basis. Since, the projection matrices $\mathbf{R}$ (on the virtual space) and the $\mathbf{P}$ (on the occupied space) are such that $\mathbf{R} + \mathbf{P} = \mathbf{1}$, we can define the final projection matrix that forms the virtual space by projecting out the core and the active orbitals given by,

$$\mathbf{R} = \mathbf{1} - \mathbf{P} \tag{8.6}$$

Thus, the projected virtual orbital $|\Psi_i\rangle$ is formed by,

$$\{|\Psi_i\rangle\} = \{\mathbf{R}|\Phi_i\rangle\} \tag{8.7}$$

where Φ_i is an atomic orbital. These virtual orbitals have redundancies or linear degeneracies as they form n orbitals spanning a space of n_{vir} orbitals. However,

a non-redundant set of virtual orbitals can be defined in terms of the linear combinations of these PAOs, and n_{vir} non-redundant virtual PAOs can be obtained by orthogonalizing the redundant virtual PAOs.

The locality of the orbitals allow us to group them into pairs of orbitals that interact strongly with each other based on their relative distances. The lower computational scaling of the local CT method arises from constraining the excitation operator A to contain only excitations from neighboring groups of orbitals,

$$\begin{aligned} A &= A_i^a(a_i^a - a_a^i) + A_{ij}^{ab}(a_{ij}^{ab} - a_{ab}^{ij}) + A_{ij}^{ak}(a_{ij}^{ak} - a_{ak}^{ij}) \\ A &= A_i^{[i]}(a_i^{[i]} - a_{[i]}^i) + A_{ij}^{[ij]}(a_{ij}^{[ij]} - a_{[ij]}^{ij}) + A_{ij}^{[ij']}(a_{ij}^{[ij']} - a_{[ij']}^{ij}) \end{aligned} \tag{8.8}$$

where $i, j, k \ldots$ denote orbitals in the active space and $a, b, c \ldots$ denote orbitals in the external space (fully occupied and virtual). The orbitals $[ij]$ denote the orbitals in the external space that form the neighboring pairs with active orbitals i or j, $[i]$ denotes the orbitals in the external space that form the neighboring pairs with active orbital i and $[ij']$ denote the orbitals in the external and active space respectively that form the neighboring pairs with active orbitals i or j. This forms the so-called *local correlation space*. Since the number of orbitals (active or external) close to ith active orbital is a constant, the scaling of the operator A can be reduced from $n_{act}^2 n_{ext}^2$ to n_{act}^2 (n_{act} and n_{ext} denote the numbers of orbitals in active and external space respectively) which is affordable even for large systems.

While very efficient implementations of this constrained excitations can be envisaged, due to the complicated nature of the CT equations, the first step towards achieving quadratic scaling local CT theory would be to partition the system completely and carry out a large number of small CT calculations. In other words the non-local CT calculation for the complete large system can be

reduced into a large number of smaller CT calculations. Say, if we have a large system and divide it into n parts $A, B, C \ldots N$, the first approximation would be to perform n CT calculations on each of the parts $A, B, C \ldots N$. Thereby the energy of system parts can be written as,

$$\begin{aligned} E_A &= \mathrm{Tr}[\Gamma_A, e^{-A_A} H_A e^{A_A}] \\ E_B &= \mathrm{Tr}[\Gamma_B, e^{-A_B} H_B e^{A_B}] \\ &\vdots \end{aligned} \tag{8.9}$$

where E_A and E_B are the energies of the system parts A and B, and H_A and H_B denote the parts of the Hamiltonian that are located on A and B respectively, and the density matrices (Γ_A, Γ_B) and amplitudes are also defined in the respective parts of the system. While this method would work for a system which can be partitioned into A and B such that there is negligible electronic interaction in between A and B, this would not be sufficient to describe real systems in which A and B interact. In such cases we would have to define E_{AB} and E_{ABC}, etc terms such that A, B and C are neighboring parts of the system which interact strongly, and E_{AB} and E_{ABC} form the energies of the combined system parts $A + B$ and $A + B + C$ respectively . Following this scheme of partitioning up to the fourth order term or E_{ABCD}, the final energy expression would look like,

$$\begin{aligned} E_{total} &= (E_A + E_B + E_C + \cdots) + (\Delta E_{AB} + \Delta E_{AC} + \cdots) \\ &\quad + (\Delta E_{ABC} + \Delta E_{ABD} + \cdots) + (\Delta E_{ABCD} + \Delta E_{ABCE} + \cdots) \end{aligned} \tag{8.10}$$

where the only contributions of ΔE_{AB}, ΔE_{ABC} and ΔE_{ABCD} are computed when A, B, C, D are neighboring parts of the system. In Eqn. 8.10, the terms ΔE_{AB}, ΔE_{ABC} and ΔE_{ABCD} refer to the corrections in contributions to the en-

ergy, given by

$$\begin{aligned}
\Delta E_{AB} &= E_{AB} - (E_A + E_B) \\
\Delta E_{ABC} &= E_{ABC} - (\Delta E_{AB} + \Delta E_{AC} + \Delta E_{BC}) - (E_A + E_B + E_C) \\
\Delta E_{ABCD} &= E_{ABCD} - (\Delta E_{ABC} + \Delta E_{ABD} + \Delta E_{BCD}) - \\
&\quad (\Delta E_{AB} + \Delta E_{AC} + \cdots) - (E_A + E_B + E_C + E_D). \qquad (8.11)
\end{aligned}$$

Since, we have used up to 4 body terms E_{ABCD} it might seem that the scaling of the local method thus formed would be $O(n^4)$. However, we are taking into account only those 2,3,4 body terms that obey the criterion that the system parts contributing to the terms are near each other and therefore, strongly interact with each other. Using this constraint on the small CT calculations performed, it can be shown geometrically that the number of CT calculations is $O(n)$, where n is the number of system parts which is proportional to the system size. Thus, we have a possible local CT method that would scale as $O(n)$ with the system size.

BIBLIOGRAPHY

[1] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Phys. Rev. Lett. **96**, 220601 (2006).

[2] D. Porras, F. Verstraete, and J. I. Cirac, Phys. Rev. B **73**, 14401 (2006).

[3] G. Vidal, Phys. Rev. Lett. **99**, 220405 (2007).

[4] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **124**, 194106 (2006).

[5] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **127**, 104107 (2007).

[6] C. Hampel and H.-J. Werner, J. Chem. Phys. **104**, 6286 (1996).

[7] M. Schutz, G. Hetzer, and H.-J. Werner, J. Chem. Phys. **111**, 5691 (1999).

[8] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

APPENDIX A

A.1 Pipek-Mezey localization scheme

The measure of delocalization defined in the Pipek-Mezey localization scheme is given by,

$$d_i = [\sum_{A=1}^{n}(Q_A^i)^2]^{-1} = [\sum_{A=1}^{n}(\sum_{\mu\in A}\sum_{\nu=1}^{m} C_{i\mu}C_{i\nu}S_{\mu\nu})^2]^{-1} \tag{A.1}$$

where Q_A^i is the gross Mulliken population of orbital i on atom A, $C_{i\mu}$ is the MO coefficient of μth AO on ith MO and $S_{\mu\nu}$ is the AO overlap between the AOs μ and ν. The number of atoms is given as n. Thus, d_i is the average number of atoms that the orbitals are delocalized on and serves as an efficient measure of delocalization. The PML scheme is to minimize this d_i.

The average of this quantity is defined as the mean localization quantity

$$D^{-1} = \frac{1}{N}\sum_{i=1}^{N} d_i^{-1}, \tag{A.2}$$

where N is the total number of molecular orbitals.

The global localization criterion is thus

$$P = \sum_{i=1}^{N}\sum_{A=1}^{n}(Q_A^i)^2 = \text{maximum}, \tag{A.3}$$

where N is the number of MOs.

This criterion is followed by a converging iteration of 2×2 rotations on the delocalized canonical orbitals as the starting orbitals to get the localized orbital

and the angle of rotation in given by,

$$\begin{aligned}
\gamma_{st} &= \text{SIGN}(B_{st})\frac{1}{4}cos^{-1}(-\frac{A_{st}}{\sqrt{A_{st}^2+B_{st}^2}}) \\
& -\frac{\pi}{4} \leq \gamma_{st} \leq \frac{\pi}{4} \\
& A_{st} = \sum_{A=1}^{n}(Q_A^{st})^2 - \frac{1}{4}[Q_A^s - Q_A^t]^2 \\
& B_{st} = \frac{1}{2}\sum_{\mu\in A}\sum_{\nu=1}^{m}[C_{s\nu}C_{t\mu} + C_{s\mu}C_{t\nu}]S_{\mu\nu}. \qquad \text{(A.4)}
\end{aligned}$$

A.2 Algorithm for parallel four-index integral transformation

Algorithm 6: Parallel four-index integral transformation algorithm.

Stage 1: Assemble J and K intermediates

Divide AO integrals $v^{AO}_{\mu\nu\kappa\lambda}$ by a factor $(2-\delta_{\mu\lambda})(2-\delta_{\nu\kappa})(2-\delta_{\mu\lambda,\nu\kappa})$

for $\bar{\nu},\bar{\kappa}\,(\bar{\nu}\geq\bar{\kappa})\in proc$ **do**

  for a,μ,λ s.t. $\mu\geq\lambda$, $\mu\lambda\geq\bar{\nu}\bar{\kappa}$ **do**

    $M^a_\mu(\bar{\nu},\bar{\kappa})$ += $v^{\mathrm{AO}}_{\mu\bar{\nu}\bar{\kappa}\lambda}\,U_{a\lambda}$; $\quad N^a_\lambda(\bar{\nu},\bar{\kappa})$ += $v^{\mathrm{AO}}_{\mu\bar{\nu}\bar{\kappa}\lambda}\,U_{a\mu}$

    $N^a_\mu(\bar{\nu},\bar{\kappa})$ += $v^{\mathrm{AO}}_{\mu\bar{\kappa}\bar{\nu}\lambda}\,U_{a\lambda}$; $\quad N^a_\lambda(\bar{\nu},\bar{\kappa})$ += $v^{\mathrm{AO}}_{\mu\bar{\kappa}\bar{\nu}\lambda}\,U_{a\mu}$

  end for

  for a,λ **do**

    $N^a_\lambda(\bar{\nu},\bar{\kappa})$ += $M^a_\lambda(\bar{\nu},\bar{\kappa})$

  end for

  for a,μ,λ s.t. $\mu\geq\lambda$, $\bar{\nu}\bar{\kappa}\geq\mu\lambda$ **do**

    $L^a_\mu(\bar{\nu},\bar{\kappa})$ += $v^{\mathrm{AO}}_{\bar{\nu}\mu\lambda\bar{\kappa}}\,U_{a\lambda}$

  end for

  for a,b,λ s.t. $a\geq b$ **do**

    $J_{ab}(\bar{\nu},\bar{\kappa})$ += $M^a_\lambda(\bar{\nu},\bar{\kappa})\,U_{b\lambda}+M^b_\lambda(\bar{\nu},\bar{\kappa})\,U_{a\lambda}+L^a_\lambda(\bar{\nu},\bar{\kappa})\,U_{b\lambda}+L^b_\lambda(\bar{\nu},\bar{\kappa})\,U_{a\lambda}$

  end for

  for a,b,λ **do**

    $K_{ab}(\bar{\nu},\lambda)$ += $N^a_\lambda(\bar{\nu},\bar{\kappa})\,U_{b\bar{\kappa}}$

  end for

end for

for a,b s.t. $a\geq b$ **do**

  write J_{ab}, K_{ab}, and K_{ba} on disk

end for

Stage 2: Redistribute J and K, transform to final integrals.
for $a, b\,(a \geq b)$ **do**
  read J_{ab}, K_{ab}, K_{ba} from disk and send to proc(a, b)
end for
for $\bar{a}, \bar{b}\,(\bar{a} \geq \bar{b}) \in$ proc, $\nu, \kappa\,(\nu \geq \kappa)$ **do**
  $J_{\bar{a}\bar{b}}(\kappa, \nu)$ += $J_{\bar{a}\bar{b}}(\nu, \kappa)$
end for
for $\bar{a}, \bar{b}\,(\bar{a} \geq \bar{b}) \in$ proc, ν, κ **do**
  $K_{\bar{a}\bar{b}}(\kappa, \nu)$ += $K_{\bar{b}\bar{a}}(\nu, \kappa)$
end for
for $\bar{a}, \bar{b}\,(\bar{a} \geq \bar{b}) \in$ proc, p, q, ν, κ **do**
  $v_{\bar{a}pq\bar{b}}$ += $J_{\bar{a}\bar{b}}(\nu, \kappa)\, U_{p\nu} U_{q\kappa}$ (eqn. (3.28))
  $v_{\bar{a}p\bar{b}q}$ += $K_{\bar{a}\bar{b}}(\nu, \kappa)\, U_{p\nu} U_{q\kappa}$ (eqn. (3.29))
end for

A.3 Algorithms for computing different elements of the two-particle density matrix from the DMRG wavefunction

In order to compute the two-particle density matrix with a reasonable cost, we use the canonical representation of the DMRG wavefunctions at different sites. By using the flexibility, the memory requirement can be reduced to $O(M^2k^2)$ and the computational cost can be reduced to $O(M^2k^4)$.

Algorithm 7: Pseudocode for COMPUTE$(4,0,0)$ function called during the first configuration of the sweep in the two-particle density matrix calculation.

for $i \in left$ **do**

  for $j \in left$ **do**

    for $k \in left$ **do**

      for $l \in left$ **do**

        Build operator $\hat{o}_i\hat{o}_j\hat{o}_k\hat{o}_l$

        $\gamma(ijkl) = parity(ijkl)\langle\Psi|\hat{o}_i\hat{o}_j\hat{o}_k\hat{o}_l \otimes \mathbf{1}_\bullet \otimes \mathbf{1}_{\mathbf{right}}|\Psi\rangle$

      end for

    end for

  end for

end for

Algorithm 8: Pseudocode for COMPUTE$(3,1,0)$ function called during the first configuration of the sweep in the two-particle density matrix calculation.

for $i \in left$ **do**

  for $j \in left$ **do**

    for $k \in left$ **do**

      Build operator $\hat{o}_i\hat{o}_j\hat{o}_k$

      for $l \in \bullet$ **do**

        $\gamma(ijkl) = parity(ijkl)\langle\Psi|\hat{o}_i\hat{o}_j\hat{o}_k \otimes \hat{o}_l \otimes \mathbf{1}_{\mathbf{right}}|\Psi\rangle$

      end for

    end for

  end for

end for

Algorithm 9: Pseudocode for COMPUTE$(3, 0, 1)$ function called during the first configuration of the sweep in the two-particle density matrix calculation.

for $i \in left$ **do**

  for $j \in left$ **do**

    for $k \in left$ **do**

      Build operator $\hat{o}_i\hat{o}_j\hat{o}_k$

      for $l \in right$ **do**

        $\gamma(ijkl) = parity(ijkl)\langle\Psi|\hat{o}_i\hat{o}_j\hat{o}_k \otimes \mathbf{1}_\bullet \otimes \hat{o}_l|\Psi\rangle$

      end for

    end for

  end for

end for

Algorithm 10: Pseudocode for COMPUTE$(2, 1, 1)$ function called during first configuration of the sweep in the two-particle density matrix calculation.

for $i \in left$ **do**

  for $j \in left$ **do**

    Load operator $\hat{o}_i\hat{o}_j$

    for $k \in \bullet$ **do**

      Load operator $\hat{o}_k$

      Build operator $\hat{o}_i\hat{o}_j\hat{o}_k$

      for $l \in right$ **do**

        $\gamma(ijkl) = parity(ijkl)\langle\Psi|\hat{o}_i\hat{o}_j \otimes \hat{o}_k \otimes \hat{o}_l|\Psi\rangle$

      end for

    end for

  end for

end for

Algorithm 11: Pseudocode for COMPUTE$(1, 2, 1)$ function called during the sweep through the block configuration in the two-particle density matrix calculation.

Initialize array of length $i \geq j \in \bullet \otimes k \in left$

for $i \in \bullet$ **do**

  for $j \leq i \in \bullet$ **do**

    Load operator $\hat{o}_i \hat{o}_j$

    for $k \in left$ **do**

      Build operator $\hat{o}_i \hat{o}_j \otimes \hat{o}_k$ and save in array

    end for

  end for

end for

for $l \in right$ **do**

  $\gamma(ijkl) = parity(ijkl) \langle \Psi | \hat{o}_i \hat{o}_j \hat{o}_k \otimes \hat{o}_l | \Psi \rangle$

end for

Algorithm 12: Pseudocode for COMPUTE$(2,1,1)$ function called during the sweep through the block configuration in the two-particle density matrix calculation.

Initialize array of length $i \geq j \in left \otimes k \in \bullet$

for $i \in left$ **do**

  for $j \leq i \in left$ **do**

    Load operator $\hat{o}_i \hat{o}_j$

    for $k \in \bullet$ **do**

      Build operator $\hat{o}_i \hat{o}_j \otimes \hat{o}_k$ and save in array

    end for

  end for

end for

for $l \in right$ **do**

  $\gamma(ijkl) = parity(ijkl) \langle \Psi | \hat{o}_i \hat{o}_j \hat{o}_k \otimes \hat{o}_l | \Psi \rangle$

end for

Algorithm 13: Pseudocode for COMPUTE$(2,0,2)$ function called during the final configuration of the sweep in the two-particle density matrix calculation.

Initialize array of length $i \geq j \in left$

for $i \in left$ **do**

  for $j \leq i \in left$ **do**

    Build operator $\hat{o}_i\hat{o}_j$ save in array

  end for

end for

for $iproc \in 1, 2 \ldots nproc$ **do**

  for $k \in right$ **do**

    for $l \leq k \in right$ **do**

      Load operator $\hat{o}_k\hat{o}_l$

      $\gamma(ijkl) = parity(ijkl)\langle\Psi|\hat{o}_i\hat{o}_j \otimes \mathbf{1}_\bullet \otimes \hat{o}_k\hat{o}_l|\Psi\rangle$

    end for

  end for

end for

Algorithm 14: Pseudocode for COMPUTE$(1, 3, 0)$ function called during the sweep through the block configuration in the two-particle density matrix calculation.

for $i \in left$ **do**

  Load operator $\hat{o}_i$

  for $j \in \bullet$ **do**

    for $k \in \bullet$ **do**

      for $l \in \bullet$ **do**

        Build operator $\hat{o}_j\hat{o}_k\hat{o}_l$

        $\gamma(ijkl) = parity(ijkl)\langle\Psi|\hat{o}_i \otimes \hat{o}_j\hat{o}_k\hat{o}_l \otimes \mathbf{1_{right}}|\Psi\rangle$

      end for

    end for

  end for

end for

Algorithm 15: Pseudocode for COMPUTE$(0, 3, 1)$ function called during the sweep through the block configuration in the two-particle density matrix calculation.

for $i \in \bullet$ **do**

  for $j \in \bullet$ **do**

    for $k \in \bullet$ **do**

      Build operator $\hat{o}_i\hat{o}_j\hat{o}_k$

      for $l \in right$ **do**

        Load operator $\hat{o}_l$

        $\gamma(ijkl) = parity(ijkl)\langle\Psi|\mathbf{1}_{\mathbf{left}} \otimes \hat{o}_i\hat{o}_j\hat{o}_k \otimes \hat{o}_l|\Psi\rangle$

      end for

    end for

  end for

end for

Algorithm 16: Pseudocode for COMPUTE$(0, 4, 0)$ function called during the sweep through the block configuration in the two-particle density matrix calculation.

for $i \in \bullet$ **do**

  for $j \in \bullet$ **do**

    for $k \in \bullet$ **do**

      for $l \in \bullet$ **do**

        Build operator $\hat{o}_i\hat{o}_j\hat{o}_k\hat{o}_l$

        $\gamma(ijkl) = parity(ijkl)\langle\Psi|\mathbf{1_{left}} \otimes \hat{o}_i\hat{o}_j\hat{o}_k\hat{o}_l \otimes \mathbf{1_{right}}|\Psi\rangle$

      end for

    end for

  end for

end for

Algorithm 17: Pseudocode for COMPUTE$(0, 0, 4)$ function called during the final configuration of the sweep in the two-particle density matrix calculation.

for $i \in right$ **do**

  for $j \in right$ **do**

    for $k \in right$ **do**

      for $l \in right$ **do**

        Build operator $\hat{o}_i\hat{o}_j\hat{o}_k\hat{o}_l$

        $\gamma(ijkl) = parity(ijkl)\langle\Psi|\mathbf{1_{left}} \otimes \mathbf{1_{right}} \otimes \hat{o}_i\hat{o}_j\hat{o}_k\hat{o}_l|\Psi\rangle$

      end for

    end for

  end for

end for